

# DIFFERENTIABILITY OF LIPSCHITZ FUNCTIONS IN LEBESGUE NULL SETS

DAVID PREISS AND GARETH SPEIGHT

**ABSTRACT.** We show that if  $n > 1$  then there exists a Lebesgue null set in  $\mathbb{R}^n$  containing a point of differentiability of each Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ ; in combination with the work of others, this completes the investigation of when the classical Rademacher theorem admits a converse. Avoidance of  $\sigma$ -porous sets, arising as irregular points of Lipschitz functions, plays a key role in the proof.

## 1. INTRODUCTION

Rademacher's theorem that Lipschitz functions on  $\mathbb{R}^n$  are differentiable almost everywhere is intrinsically important and has been the source of many modern developments. From the developments that are not directly related to the present work, we feel we just have to mention, at least briefly, the work of Cheeger [4], Keith [11] and Bate [2], which starting from the notion of metric measure spaces satisfying the Poincaré inequality eventually led to understanding differentiability of Lipschitz functions on metric spaces in a way similar to Rademacher's theorem. On the other hand, the investigation of validity of an infinite dimensional generalization of Rademacher's theorem, a survey of older results in Chapters 4–6 of the authoritative book by Benyamini and Lindenstrauss [3], and very recent progress presented in the recent research monograph by Lindenstrauss, Preiss, and Tišer [14] has been basic for much of what we do here.

To introduce our result, the most natural formulation of the classical Rademacher theorem states that if a Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is differentiable at no point of a set  $A \subset \mathbb{R}^n$ , then  $A$  must be Lebesgue null. The natural converse of this statement asks: given a Lebesgue null set  $A \subset \mathbb{R}^n$ , does there exist a Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  which is differentiable at no point of  $A$ ? The answer has been long known to be positive and relatively easy in the case  $m = n = 1$ , when the statement of Rademacher's theorem is a special case of Lebesgue's differentiation of monotone functions. Although we are unable to find the first reference to this result, we may refer the reader to the full description of sets of non-differentiability of real valued Lipschitz functions on the real line due to Zahorski [19], or to a modern variant of Zahorski's argument in [8].

---

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.2011-ADG-20110209. The second author was supported by EPSRC funding.

The discovery that the converse to the higher dimensional version of Rademacher's theorem is not straightforward came originally as a byproduct of an infinite dimensional differentiability result of Preiss [15]: for  $n > 1$  there is a Lebesgue null set in  $\mathbb{R}^n$  containing a point of differentiability for every real valued Lipschitz function: for example, any Lebesgue null  $G_\delta$  set containing all lines passing through distinct points with rational coordinates has this property. It seems probable that, similarly to what happened with [15], a modification of the proof of [14] of a differentiability result for Lipschitz maps of infinite dimensional Hilbert spaces to  $\mathbb{R}^2$  would lead to showing that the converse to Rademacher's theorem fails for maps from  $\mathbb{R}^n$  to  $\mathbb{R}^2$  for  $n > 2$ . However, these authors also show that the corresponding differentiability result for  $\mathbb{R}^3$  valued maps is false, thereby indicating that for general  $m, n$  new methods are needed.

In combination with two recently announced developments, by Alberti, Csörnyei, and Preiss [1] and by Csörnyei and Jones [5], our result completely answers the question of validity of the converse to the Rademacher theorem: it holds if and only if  $m \geq n$ . Indeed, [1] shows this when  $n = 2$  and, for general  $n$ , provides necessary and sufficient geometric criteria for a set to be contained in the non-differentiability set of a Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . The sets satisfying these criteria form a  $\sigma$ -ideal, implying that for any given  $n$  the problem of validity of the converse to the Rademacher theorem for functions  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  has the same answer for  $m \geq n$ . The question whether this  $\sigma$ -ideal coincides with the  $\sigma$ -ideal of Lebesgue null sets was open until Csörnyei and Jones [5] announced a very deep and difficult result showing that this is indeed the case. Together, these results imply that the converse to the Rademacher theorem is true provided that  $m \geq n$ . Here we fill in the last piece of the puzzle by showing that in all remaining cases the converse actually fails.

**Theorem 1.1.** *Suppose  $n > 1$ . Then there exists a Lebesgue null set  $N \subset \mathbb{R}^n$  containing a point of differentiability for every Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ .*

Perhaps surprisingly, our approach to proving this is not related to known results on  $\varepsilon$ -differentiability. The notion of  $\varepsilon$ -differentiability is defined similarly to differentiability but with fixed error  $\varepsilon$  in the first order approximation of a function by its derivative (rather than arbitrarily small error on sufficiently small scales). The  $\varepsilon$ -differentiability results appeared first in the infinite dimensional context in [12] (see Chapter 4 of [14] for further developments). The restriction of their proof to the finite dimensional situation would provide a Lebesgue null set  $N \subset \mathbb{R}^n$  such that every Lipschitz map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-2}$  has points of  $\varepsilon$ -differentiability inside  $N$  for every  $\varepsilon > 0$ . This was improved by De Pauw and Huovinen [6]: for  $n > 1$  there exists a Lebesgue null set  $N \subset \mathbb{R}^n$  such that every Lipschitz map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  has points of  $\varepsilon$ -differentiability inside  $N$  for every  $\varepsilon > 0$ . The reason why the present development does not build on  $\varepsilon$ -differentiability results is treated in detail in the infinite dimensional context throughout [14]. Here we just mention that, while  $\varepsilon$ -differentiability has been proved for Lipschitz maps of a Hilbert space to any  $\mathbb{R}^n$ , differentiability is known only for Lipschitz maps to  $\mathbb{R}$  or  $\mathbb{R}^2$ .

The direction our approach to proving Theorem 1.1 took started with the second named author's answer in [16] to the question from [14] on size of  $\sigma$ -porous sets. These sets (defined in Definition 4.4) form a subclass of non-differentiability sets, and one of the contributions of [14] was in better understanding of what was noticed in [13], that knowledge of smallness of porous sets is an important step in proving a differentiability result. Results of Chapter 10 in [14] show that in many spaces, including the finite dimensional ones,  $\sigma$ -porous sets are null on typical curves as well as on typical 2-dimensional surfaces (where 'typical' is understood in the sense of Baire category). The result of Speight [16] shows that even in  $\mathbb{R}^4$  this is no longer the case for 3-dimensional surfaces. This opened the door to questions whether the deep infinite dimensional counterexamples from Chapter 14 of [14] have an analogy in the finite dimensional situation. By discovering that the answer is no, the second named author made an important, and as it turned out decisive, step toward the proof of Theorem 1.1.

It is natural to ask how small can be sets  $N \subset \mathbb{R}^n$  inside which one may find a point of differentiability of every Lipschitz  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ . For  $n > m = 1$ , this was studied by Doré and Maleva [4] who found such sets can be made compact and of Hausdorff dimension one, and recently Dymond and Maleva announced that they can make them of Minkowski dimension one. However, it is not clear whether such sets can be purely unrectifiable (meeting every rectifiable curve in a set of one dimensional measure zero). It seems possible that the finite dimensional analogy of the above mentioned  $\mathbb{R}^2$  valued differentiability result of [14] may lead to similar improvements for  $n > m = 2$ . However, our method differs significantly from the methods used to prove these improvements; so at the present time we can only notice that a simple modification of our arguments (explained in Remark 9.2) provides, for any  $n > m$  and  $\tau > 0$ , a set  $N \subset \mathbb{R}^n$  of Hausdorff dimension  $m + \tau$  containing a point of differentiability of every Lipschitz  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

We now fix  $n > 1$  for the remainder of the paper and briefly describe the structure of the proof of Theorem 1.1. It follows that of a proof by Lindenstrauss, Preiss, and Tišer [14, Theorem 13.1.1], which constructs points of Fréchet differentiability for vector valued Lipschitz functions on infinite dimensional Banach spaces satisfying particular smoothness assumptions.

The idea behind both of the proofs is to apply a variational principle (Lemma 6.1) so that a perturbation of the function

$$(x, T) \mapsto \|f'(x; T)\|_H^2$$

attains a maximum at some  $(x_\infty, T_\infty)$ . Here  $(x, T)$  are pairs such that  $f$  is regularly differentiable at  $x$  in the direction  $T$  (Definition 4.1),  $f'(x; T)$  is the directional derivative of  $f$  at  $x$  in direction  $T$ , and  $\|\cdot\|_H$  is the Hilbert-Schmidt norm. Intuitively,  $f$  is regularly differentiable at  $x$  in direction  $T$  if changes in values of  $f$  are approximated by  $f'(x; T)$  not only on planes with direction  $T$  passing through  $x$  but also on nearby parallel planes.

One assumes  $f$  is not differentiable at  $x_\infty$  and uses regular differentiability to find a pair  $(x, T)$  such that  $f$  is regularly differentiable at  $x$  in direction  $T$  and  $\|f'(x; T)\|_H^2$  is larger than  $\|f'(x_\infty, T_\infty)\|_H^2$ . Since only a perturbation of the directional derivative was maximized, to get a contradiction, one

must also choose  $(x, T)$  so that the corresponding change in the perturbing functions is small relative to  $\|f'(x; T)\|_H^2 - \|f'(x_\infty; T_\infty)\|_H^2$ .

To prove Theorem 1.1 we consider points  $x$  in a Lebesgue null set and need to find points of regular differentiability in the direction of  $n - 1$  dimensional planes. To do this we use the fact that points where  $f$  is differentiable, but not regularly differentiable, in some direction are irregular (Definition 4.2 and Lemma 4.3). Further, the irregular points of  $f$  form a  $\sigma$ -porous set (Lemma 4.6).

Intuitively, a set is porous (Definition 4.4) if each point of the set sees relatively large holes in the set on arbitrarily small scales. A set is  $\sigma$ -porous if it is a countable union of porous sets. The collection of  $\sigma$ -porous sets in  $\mathbb{R}^n$  is strictly contained in the family of meager Lebesgue null sets. For a survey of porous sets, including their interesting applications to differentiability, see [20] and [21].

It follows from work of Tukia [17] that there exists a doubling measure (Definition 2.1)  $\mu$  on  $\mathbb{R}$  which gives full measure to a Lebesgue null set  $\tilde{N} \subset \mathbb{R}$  (Theorem 5.1). Since porous sets have measure zero with respect to doubling measures (Proposition 4.5) and  $\mathcal{L}^{n-1} \times \mu$  is doubling, we can find many points in the Lebesgue null set  $\mathbb{R}^{n-1} \times \tilde{N}$  outside a given  $\sigma$ -porous set. The Lebesgue null set in Theorem 1.1 is constructed using affine copies of the set  $\mathbb{R}^{n-1} \times \tilde{N}$  (Definition 5.3). Since irregular points of  $f$  form a  $\sigma$ -porous set, it follows that such a set contains many points of regular differentiability of  $f$  in the direction of  $n - 1$  dimensional planes. The fact that the measure  $\mu$  is doubling also helps us to control the perturbation terms.

## 2. DOUBLING MEASURES

We first recall, and for completeness prove, some basic facts about doubling measures.

**Definition 2.1.** A Borel measure  $\mu$  on a metric space  $M$  is *doubling* if balls have finite positive measure and there exists  $C \geq 1$  such that

$$\mu(B(x, 2r)) \leq C\mu(B(x, r))$$

for all  $x \in M$  and  $r > 0$ .

Doubling measures give relatively large measure to subballs of relatively large radius.

**Lemma 2.2.** Suppose  $\mu$  is a doubling measure on  $\mathbb{R}^m$ . Then there exists a constant  $C(\mu) \geq 1$  such that

$$\frac{\mu(B(y, s))}{\mu(B(x, r))} \leq C(\mu) \left(\frac{s}{r}\right)^{C(\mu)}$$

whenever  $B(x, r) \subset B(y, s)$ .

*Proof.* Let  $N$  be an integer such that  $\log_2(s/r) + 1 \leq N \leq \log_2(s/r) + 2$ . Then  $2^N r \geq 2s$  and so  $B(y, s) \subset B(x, 2s) \subset B(x, 2^N r)$ . Hence,

$$\frac{\mu(B(y, s))}{\mu(B(x, r))} \leq \frac{\mu(B(x, 2^N r))}{\mu(B(x, r))} \leq C(\mu)^N \leq C(\mu)^{\log_2(s/r)+2} \leq C(\mu) \left(\frac{s}{r}\right)^{C(\mu)}.$$

□

We also need to know that doubling measures give small measures to thin annuli, independently of the centre and radius of the associated ball.

**Proposition 2.3.** *Suppose  $\mu$  is a doubling measure on  $\mathbb{R}^m$  and  $\varepsilon > 0$ . Then there is  $\delta > 0$  such that*

$$\frac{\mu(B(x, (1-t)r))}{\mu(B(x, r))} \geq 1 - \varepsilon$$

whenever  $x \in \mathbb{R}^m$ ,  $r > 0$  and  $0 < t < \delta$ .

*Proof.* Suppose  $x \in \mathbb{R}^m$  and  $r > 0$  with

$$\mu(B(x, r) \setminus B(x, (1-2^{-N})r)) > \alpha\mu(B(x, r))$$

for some  $N \in \mathbb{N}$  and  $\alpha > 0$ . Define, for  $i = 1, \dots, N$ ,

$$A_i = B(x, (1-2^{-i-1})r) \setminus B(x, (1-2^{-i})r).$$

Fix  $1 \leq i \leq N$ . Choose a (necessarily finite) collection  $B(x_j, 2^{-i-2}) \subset A_i$  of disjoint open balls such that, for a constant  $C_1 > 0$  depending only on  $m$ ,

$$B(x, r) \setminus B(x, (1-2^{-N})r) \subset \bigcup_j B(x_j, C_1 2^{-i-2}).$$

Since the balls  $B(x_j, 2^{-i-2})$  are disjoint it follows from Lemma 2.2 that, for some constant  $C$  depending on  $C_1$  and the doubling constant of  $\mu$ ,

$$\mu(B(x, r) \setminus B(x, (1-2^{-N})r)) \leq C \sum_j \mu(B(x_j, 2^{-i-1})) \leq C\mu(A_i).$$

Since the sets  $A_i$  are disjoint we deduce,

$$\mu(B(x, r)) \geq \sum_{i=1}^N \mu(A_i) \geq N\alpha\mu(B(x, r))/C.$$

Hence by choosing  $N$  sufficiently large (independently of  $x$  and  $r$ ) we can ensure  $\alpha$  is small. This gives the desired conclusion.  $\square$

**Definition 2.4.** If  $\mu$  is a doubling measure on  $\mathbb{R}^m$  and  $f: \mathbb{R}^m \rightarrow \mathbb{R}$  is locally integrable we define the *maximal operator*  $M_\mu f$  by

$$M_\mu f(x) = \sup_{r>0} \frac{1}{\mu(B(x, r))} \int_{B(x, r)} |f|.$$

We recall that if  $\mu$  is doubling then the maximal operator cannot increase the norm on  $L^2(\mu)$  too much [9, Theorem 2.2].

**Lemma 2.5.** *Suppose  $\mu$  is a doubling measure on  $\mathbb{R}^m$  and  $f \in L^2(\mu)$ . Then there is a constant  $C(\mu) \geq 1$  such that*

$$\|M_\mu f\|_{L^2(\mu)} \leq C(\mu) \|f\|_{L^2(\mu)}.$$

### 3. REGULARITY DEFECT

Recall that a function  $f: M \rightarrow N$  between metric spaces is *Lipschitz* if there is a constant  $L \geq 0$  such that

$$d(f(x), f(y)) \leq Ld(x, y)$$

for all  $x, y \in M$ . We denote the smallest such  $L$ , called the *Lipschitz constant* of  $f$ , by  $\text{Lip}(f)$ .

We now discuss regularity and derive estimates needed for the proof of Theorem 1.1. Note we identify  $\mathbb{R}^{n-1}$  with the subspace of  $\mathbb{R}^n$  in which the final coordinate is zero.

**Definition 3.1.** Let  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  be a Lipschitz function and  $u \in \mathbb{R}^n$ . We define the *defect of regularity* of  $f$  at the point  $u$  by

$$\text{reg}_{n-1}f(u) = \sup_{\substack{|v|+|w|>0 \\ v-w \in \mathbb{R}^{n-1}}} \frac{|f(u+v) - f(u+w)|}{|v| + |w|}.$$

If  $\Omega \subset \mathbb{R}^n$  we define  $\text{reg}_{n-1}f(u, \Omega)$ , the *defect of restricted regularity*, by the same formula but with the requirement that the open straight segment  $(u+v, u+w)$  lies in the set  $\Omega$ .

Intuitively, the regularity defect of  $f$  at  $u$  describes how much  $f$  varies on line segments, with direction in  $\mathbb{R}^{n-1}$ , which are not too short compared to their distance from  $u$ .

If  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  we denote the derivative of  $f$  with respect to the first  $n-1$  variables by  $f'_{n-1}$ . Thus  $f'_{n-1}(u) \in L(\mathbb{R}^{n-1}, \mathbb{R}^{n-1})$  whenever it exists.

The proof of the following lemma is adapted from the proof of a similar result by Lindenstrauss, Preiss, and Tišer [14, Lemma 9.5.4, Lemma 13.2.3]. The main difference is that in one direction we have a general doubling measure rather than Lebesgue measure.

**Lemma 3.2.** Let  $\mu$  be a doubling measure on  $\mathbb{R}$  and  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  be Lipschitz. Suppose  $a \in \mathbb{R}^n$  and  $R > 0$ . Then there is a constant  $C(\mu) \geq 1$  (independent of  $f$ ,  $a$ , and  $R$ ) such that

$$\begin{aligned} & (\text{reg}_{n-1}f(a, B(a, R)))^{n+C(\mu)-1} \\ & \leq C(\mu)(\text{Lip}(f))^{n+C(\mu)-2} M_{\mathcal{L}^{n-1} \times \mu}(1_{B(a, R)} \|f'_{n-1}\|)(a). \end{aligned}$$

*Proof.* Since the statement is translation invariant and homogeneous with respect to  $f$  we can assume  $a = 0$  and  $\text{Lip}(f) = 1$ . Fix  $u, v \in B(0, R)$  with  $u - v \in \mathbb{R}^{n-1}$ . Suppose  $r = |f(u) - f(v)| > 0$  and let  $S = \max(|u|, |v|)$ . Define  $\tau = |u - v|$  and  $e = (u - v)/\tau$ .

There exists a constant  $C$  and  $x, y \in \mathbb{R}^n$  with  $x - y \in \mathbb{R}^{n-1}$  such that if  $z \in B(x, Cr)$  then  $z + te \in B(0, S)$  for all  $0 \leq t \leq \tau$ ,  $z + \tau e \in B(y, Cr)$  and  $|f(z + \tau e) - f(z)| \geq Cr$ . Hence, for each  $z \in B(x, Cr)$  (not distinguishing between constants  $C$  and  $1/C$ ),

$$r \leq C \left| \int_0^\tau \frac{d}{dt} f(z + te) dt \right| \leq C \int_0^\tau \|f'_{n-1}(z + te)\| dt.$$

By integrating over  $z \in B(x, Cr)$  with  $\langle z, e \rangle = 0$  we obtain

$$r^n \mu(B(x, Cr)) \leq C \int_{B(0, S)} \|f'_{n-1}\| d(\mathcal{L}^{n-1} \times \mu).$$

We note, since  $\mu$  is doubling,

$$\frac{\mu(B(0, S))}{\mu(B(x, Cr))} \leq C(\mu)(S/r)^{C(\mu)}.$$

The previous two inequalities imply

$$(r/S)^{n+C(\mu)-1} \leq \frac{C(\mu)}{\mathcal{L}^{n-1} \times \mu(B(0, S))} \int_{B(0, S)} \|f'_{n-1}\| \, d(\mathcal{L}^{n-1} \times \mu).$$

Hence

$$\left( \frac{|f(u) - f(v)|}{|u| + |v|} \right)^{n+C(\mu)-1} \leq C(\mu) M_{\mathcal{L}^{n-1} \times \mu}(1_{B(0, S)} \|f'_{n-1}\|)(0).$$

Finally we conclude

$$(\text{reg}_{n-1} f(0, B(0, R)))^{n+C(\mu)-1} \leq C(\mu) M_{\mathcal{L}^{n-1} \times \mu}(1_{B(0, R)} \|f'_{n-1}\|)(0)$$

as required.  $\square$

The proof of the following proposition is essentially the same as that of a similar result by Lindenstrauss, Preiss, and Tišer [14, Corollary 13.2.4].

**Proposition 3.3.** *Let  $\Omega \subset \mathbb{R}^n$  be a bounded Borel measurable set,  $\mu$  be a doubling measure on  $\mathbb{R}$ ,  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  be Lipschitz and  $s, \lambda > 0$ . Then*

$$\begin{aligned} & \mathcal{L}^{n-1} \times \mu\{u \in \mathbb{R}^n : B(u, s) \subset \Omega, \text{reg}_{n-1} f(u, B(u, s)) > \lambda\} \\ & \leq \frac{C(\mu)(\text{Lip}(f))^{2(n+C(\mu))-4}}{\lambda^{2(n+C(\mu))-2}} \int_{\Omega} \|f'_{n-1}\|^2 \, d(\mathcal{L}^{n-1} \times \mu). \end{aligned}$$

*Proof.* For each  $u \in \mathbb{R}^n$  such that  $B(u, s) \subset \Omega$  it follows, using Lemma 3.2,

$$\begin{aligned} & (\text{reg}_{n-1} f(u, B(u, s)))^{n+C(\mu)-1} \\ & \leq C(\mu)(\text{Lip}(f))^{n+C(\mu)-2} M_{\mathcal{L}^{n-1} \times \mu}(1_{\Omega} \|f'_{n-1}\|)(u). \end{aligned}$$

Using the maximal operator and Chebyshev's inequalities give,

$$\begin{aligned} & \mathcal{L}^{n-1} \times \mu\{u \in \mathbb{R}^n : B(u, s) \subset \Omega, \text{reg}_{n-1} f(u, B(u, s)) > \lambda\} \\ & \leq \frac{C(\mu)(\text{Lip}(f))^{2(n+C(\mu))-4}}{\lambda^{2(n+C(\mu))-2}} \int_{\mathbb{R}^n} M_{\mathcal{L}^{n-1} \times \mu}(1_{\Omega} \|f'_{n-1}\|)^2 \, d(\mathcal{L}^{n-1} \times \mu) \\ & \leq \frac{C(\mu)(\text{Lip}(f))^{2(n+C(\mu))-4}}{\lambda^{2(n+C(\mu))-2}} \int_{\Omega} \|f'_{n-1}\|^2 \, d(\mathcal{L}^{n-1} \times \mu). \end{aligned}$$

$\square$

From now on, since increasing the constants weakens the statements, we may assume the constants  $C(\mu)$  appearing in statements of results in this section are the same.

#### 4. REGULAR DIFFERENTIABILITY

When proving Theorem 1.1 we will work with functions differentiable only in the direction of certain subspaces. A useful stronger condition is the notion of regular differentiability in a particular direction.

**Definition 4.1.** Suppose  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $T \in L(\mathbb{R}^p, \mathbb{R}^n)$  and  $x \in \mathbb{R}^n$ .

- We say the function  $f$  is *differentiable at  $x$  in the direction of  $T$*  if there is  $f'(x; T) \in L(\mathbb{R}^p, \mathbb{R}^m)$  such that for every  $\varepsilon > 0$  there is  $\delta > 0$  such that

$$|f(x + Tv) - f(x) - f'(x; T)(v)| \leq \varepsilon|v|$$

whenever  $v \in \mathbb{R}^p$  and  $|v| < \delta$ .

- We say  $f$  is *regularly differentiable at  $x$  in the direction of  $T$*  if  $f$  is differentiable at  $x$  in the direction of  $T$  and for every  $\varepsilon > 0$  there is  $\delta > 0$  such that

$$|f(x + z + Tv) - f(x + z) - f'(x; T)(v)| \leq \varepsilon(|v| + |z|)$$

whenever  $v \in \mathbb{R}^p$ ,  $z \in \mathbb{R}^n$  and  $|v| + |z| < \delta$ .

**Definition 4.2.** Suppose  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ . We say that  $x$  is a *regular point* of  $f$  if for every  $v \in \mathbb{R}^n$  for which the directional derivative  $f'(x; v)$  exists,

$$\lim_{t \rightarrow 0} \frac{f(x + tz + tv) - f(x + tz)}{t} = f'(x; v)$$

uniformly for  $z \in \mathbb{R}^n$  such that  $|z| \leq 1$ . A point which is not regular is called *irregular*.

Intuitively regular differentiability of  $f$  at  $x$  in direction  $T$  means the change in  $f$  along lines in a direction  $Tv$  which pass close to  $x$  is well approximated by the directional derivative. Thus the following lemma can be expected.

**Lemma 4.3.** Suppose the map  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is Lipschitz,  $T \in L(\mathbb{R}^p, \mathbb{R}^m)$  and  $x \in \mathbb{R}^n$ . Suppose  $f$  is differentiable at  $x$  in the direction of  $T$  and  $x$  is a regular point of  $f$ . Then  $f$  is regularly differentiable at  $x$  in the direction of  $T$ .

*Proof.* Because  $f$  is differentiable at  $x$  in the direction of  $T$  we have that  $f'(x; Tv)$  exists for every  $v \in V$  and  $f'(x; Tv) = f'(x; T)(v)$ .

Let  $\varepsilon > 0$ . Since  $x$  is a regular point of  $f$  and the sphere in  $\mathbb{R}^p$  is compact, there exists  $\delta > 0$  such that whenever  $|v| = \varepsilon$ ,  $0 < t < \delta$  and  $|z| \leq 1$ ,

$$|f(x + tz + Ttv) - f(x + tz) - f'(x; T)(tv)| < \varepsilon^2 t.$$

Set  $w = tv$  so  $|w| = t\varepsilon$ . Then, provided  $|w| < \varepsilon\delta$  and  $|z| \leq 1$ ,

$$|f(x + \varepsilon^{-1}|w|z + Tw) - f(x + \varepsilon^{-1}|w|z) - f'(x; T)(w)| < \varepsilon|w|.$$

That is, provided  $|w| < \varepsilon\delta$  and  $|p| < \varepsilon^{-1}|w|$ ,

$$|f(x + p + Tw) - f(x + p) - f'(x; T)(w)| < \varepsilon|w|.$$

If  $|p| > \varepsilon^{-1}|w|$  then, since  $f$  is Lipschitz,

$$|f(x + p + Tw) - f(x + p) - f'(x; T)(w)| \leq 2\text{Lip}(f)\|T\||w| < 2\text{Lip}(f)\|T\|\varepsilon|p|.$$

We have shown if  $|w| < \varepsilon\delta$  then, for all  $p \in \mathbb{R}^n$ ,

$$|f(x + p + Tw) - f(x + p) - f'(x; T)(w)| < \varepsilon|w| + 2\text{Lip}(f)\|T\|\varepsilon|p|.$$

Hence  $f$  is regularly differentiable at  $x$  in the direction  $T$ .  $\square$

Intuitively a set  $P$  is porous if each point of  $P$  sees nearby, on arbitrarily small scales, relatively large holes in  $P$ .



**Definition 4.4.** A set  $P$  in a metric space  $M$  is *porous* if for each  $x \in P$  there is  $\lambda > 0$  and  $x_k \rightarrow x$  such that  $B(x_k, \lambda d(x_k, x)) \cap P = \emptyset$ . A set is  *$\sigma$ -porous* if it is a countable union of porous sets.

The following fact is well known [21]; it follows directly from the fact that a Lebesgue density theorem holds for doubling measures.

**Proposition 4.5.** *Doubling measures give measure zero to porous sets.*

The following lemma [14, Proposition 6.2.6] will be essential when finding regular points of  $f$  inside a Lebesgue null set.

**Lemma 4.6.** *Suppose  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is Lipschitz. Then the set of irregular points of  $f$  is  $\sigma$ -porous.*

## 5. DEFINITION OF THE LEBESGUE NULL SET

As already mentioned, we need our Lebesgue null set to contain plenty of regular points of Lipschitz functions. The following theorem follows from results of Tukia [17] on quasisymmetric mappings.

**Theorem 5.1.** *There exists a doubling measure on  $\mathbb{R}$  which gives full measure to a Lebesgue null set.*

Recall that a set in a metric space is of class  $G_\delta$  if it is a countable intersection of open sets. In the proof of Theorem 1.1 it will be convenient to work with a  $G_\delta$  set. The reason for this is the following theorem of Mazurkiewicz [7, Theorem 8.3].

**Lemma 5.2.** *Let  $Z$  be a complete metric space. Then  $A \subset Z$  is topologically complete if and only if it is a  $G_\delta$  set in  $Z$ .*

It follows from the definitions that every Lebesgue null set is contained in a Lebesgue null set of class  $G_\delta$ .

Let  $\mathcal{A}$  be a family of affine maps of the form  $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$  given by  $Tx = Ax + v$ , where  $A \in L(\mathbb{R}^n, \mathbb{R}^n)$  has a matrix representation (with respect to the standard basis of  $\mathbb{R}^n$ ) with rational entries and  $v \in \mathbb{Q}^n$ .

We now define the Lebesgue null set we will work with.

**Definition 5.3.** Fix a doubling measure  $\mu$  on  $\mathbb{R}$  which gives full measure to a Lebesgue null set  $\tilde{N} \subset \mathbb{R}$ . Let  $N$  be any Lebesgue null set in  $\mathbb{R}^n$  of class  $G_\delta$  containing

$$\bigcup_{A \in \mathcal{A}} A(\mathbb{R}^{n-1} \times \tilde{N}).$$

To prove Theorem 1.1 we show every Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  has a point of differentiability inside  $N$ .

## 6. SETUP OF THE VARIATIONAL PRINCIPLE

In this section we define a suitable space and perturbations then show the following variational principle [14, Corollary 7.2.4] is applicable.

**Lemma 6.1.** *Suppose that  $f: M \rightarrow \mathbb{R}$  is lower bounded and lower semicontinuous on a complete metric space  $(M, d)$ . Suppose further that functions  $F_i: M \times M \rightarrow [0, \infty]$ ,  $i \geq 0$ , are lower semicontinuous in the second variable with  $F_i(x, x) = 0$  for all  $x \in M$  and that  $0 < r_i \leq \infty$  are such that  $r_i \rightarrow 0$  and*

$$\inf_{d(x,y) > r_i} F_i(x, y) > 0.$$

*If  $x_0 \in M$  and  $(\varepsilon_i)_{i=0}^\infty$  is any sequence of positive numbers such that*

$$f(x_0) < \varepsilon_0 + \inf_{x \in M} f(x) \text{ and } \inf_{d(x_0, y) > r_0} F_0(x_0, y) > \varepsilon_0,$$

*then one may find a sequence  $(x_i)_{i=1}^\infty$  of points in  $M$  converging to some  $x_\infty \in M$  such that the function*

$$h(x) = f(x) + \sum_{i=0}^\infty F_i(x_i, x)$$

*attains its minimum on  $M$  at  $x_\infty$ . Moreover, for each  $i \geq 0$ ,*

$$d(x_i, x_\infty) \leq r_i, \quad F_i(x_i, x_\infty) \leq \varepsilon_i$$

$$h(x_\infty) \leq \varepsilon_i + \inf_{x \in M} \left( f(x) + \sum_{j=0}^{i-1} F_j(x_j, x) \right).$$

From now on we intend to make the following assumptions:

**Assumptions 6.2.** Suppose  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  is a bounded Lipschitz function and that  $0 < \eta_0 < 1/2$ ,  $x_0 \in N$  and  $T_0 \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  are such that  $\|T_0\| = 1/2$ ,  $f$  is regularly differentiable at  $x_0$  in the direction  $T_0$  and

$$\|\text{Id} - f'(x, T)\| \leq \frac{1}{4}$$

whenever  $\|T - T_0\| \leq \eta_0$  and  $f'(x; T)$  exists.

We now show it is sufficient to prove the following proposition.

**Proposition 6.3.** *Suppose  $f$  satisfies the assumptions above. Then there is a point  $x \in N$  such that  $|x - x_0| \leq \eta_0$  and  $f$  is differentiable at  $x$ .*

**Claim 6.4.** *Theorem 1.1 follows from Proposition 6.3*

*Proof.* Suppose  $f: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  is Lipschitz. By Lemma 4.5 and Lemma 4.6 the set of regular points of  $f$  has full measure with respect to  $\mathcal{L}^{n-1} \times \mu$ . Hence we may find  $t_0 \in \tilde{N}$  such that for  $\mathcal{L}^{n-1}$  almost every  $z \in \mathbb{R}^{n-1}$  the function  $f$  is regular at  $(z, t_0)$ . Using the classical Rademacher theorem for the Lipschitz function  $z \mapsto f(z, t_0)$  and Lemma 4.3, we may find  $z_0 \in \mathbb{R}^{n-1}$  and  $T_0 \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  such that  $\text{rank } T_0 = n - 1$ ,  $\|T_0\| = 1/2$  and  $f$  is regularly differentiable at  $x_0 = (z_0, t_0)$  in the direction  $T_0$ .

Let  $g$  be an extension of  $f$  restricted to  $B(x_0, 1)$  to a bounded Lipschitz function on  $\mathbb{R}^n$ . Let  $R_0 \in L(\mathbb{R}^n, \mathbb{R}^{n-1})$  be such that  $R_0 T_0 = \text{Id}$  on  $\mathbb{R}^{n-1}$ .

Fix  $\eta_0 > 0$  to be chosen small later. Consider the function  $h = R_0 + \eta_0 g$ . Assume that  $\|T - T_0\| \leq \eta_0$ . We estimate,

$$\begin{aligned} \|\text{Id} - h'(x; T)\| &= \|R_0(T_0 - T) - \eta_0 g'(x; T)\| \\ &\leq \|R_0\| \|T - T_0\| + \eta_0 \text{Lip}(g) \|T\| \\ &\leq \|R_0\| \|T - T_0\| + \eta_0 \text{Lip}(g) (\|T_0\| + \|T - T_0\|) \\ &\leq (\|R_0\| + 2\text{Lip}(g)) \eta_0 \\ &\leq \frac{1}{4} \end{aligned}$$

provided  $\eta_0$  is sufficiently small. Hence there exists  $x \in N$  with  $|x - x_0| \leq \eta_0$  at which  $h$  is differentiable. Consequently  $g$  and hence  $f$  are differentiable at  $x \in N$ , as required.  $\square$

From now on we make the assumptions given above and focus on proving Proposition 6.3. We now establish basic consequences of our assumptions. First we state a mean value estimate [14, Proposition 2.4.1].

**Lemma 6.5.** *Let  $\Lambda$  be a real valued locally Lipschitz function on an open subset  $G$  of a separable Banach space  $X$ , and let  $a, b \in G$  be such that the straight segment from  $a$  to  $b$  is contained in  $G$ . Then for every  $\varepsilon > 0$  there is a point  $z \in G$  at which  $\Lambda$  is Gâteaux differentiable and*

$$\Lambda'(z)(b - a) > \Lambda(b) - \Lambda(a) - \varepsilon.$$

**Proposition 6.6.** *The following facts hold:*

- (1) *The inequality  $|f(x + Tu) - f(x)| \leq 5|u|/4$  holds for every  $x \in \mathbb{R}^n$ , linear map  $T$  such that  $\|T - T_0\| \leq \eta_0$ , and  $u \in \mathbb{R}^{n-1}$ .*
- (2) *If  $\|T - T_0\| \leq \eta_0$  and  $f'(x; T)$  exists then  $T$  is a linear isomorphism of  $\mathbb{R}^{n-1}$  onto its image and  $f'(x; T)$  is a linear isomorphism of  $\mathbb{R}^{n-1}$  onto itself.*
- (3) *Necessarily  $\text{Lip}(f) \geq 1$ .*

*Proof.* Temporarily denote  $v = (f(x + Tu) - f(x)) / |f(x + Tu) - f(x)| \in \mathbb{R}^{n-1}$ . The map  $z \mapsto v \cdot f(x + Tz)$  is a real valued Lipschitz map on  $\mathbb{R}^{n-1}$ . Hence, by Lemma 6.5 with  $a = 0$  and  $b = u$ , for every  $\varepsilon > 0$  there is a point  $z \in \mathbb{R}^{n-1}$  at which this map is differentiable and

$$v \cdot (f'(x + Tz; T)u) > v \cdot (f(x + Tu) - f(x)) - \varepsilon.$$

Hence

$$|f(x + Tu) - f(x)| \leq \sup\{\|f'(y; T)\| : y \in \mathbb{R}^n, f'(y; T) \text{ exists}\} |u| \leq \frac{5}{4}|u|$$

which proves (1).

Suppose  $\|T - T_0\| \leq \eta_0$  and  $f'(x; T)$  exists. Then by our assumptions

$$\|\text{Id} - f'(x; T)\| \leq \frac{1}{4}.$$

It follows immediately from this that  $f'(x; T)$  is a linear injection from  $\mathbb{R}^{n-1}$  to itself and consequently is a linear isomorphism. Using the definition of  $f'(x; T)$  it follows that  $T$  must have  $n - 1$  dimensional image, hence is a linear isomorphism of  $\mathbb{R}^{n-1}$  onto its image. Hence (2) is true.

If  $u$  is any unit vector in  $\mathbb{R}^{n-1}$  then  $2\|T_0 u\| \leq 1$ . Hence

$$\text{Lip}(f) \geq 2\|f'(x_0; T_0 u)\| \geq 2(|u| - |u - f'(x; T_0 u)|) \geq 1.$$

This proves (3).  $\square$

Recall that if  $H$  is a Hilbert space then the *Hilbert-Schmidt norm*  $\|\cdot\|_H$  and corresponding inner product  $\langle \cdot, \cdot \rangle_H$  are defined on the space  $L(H, \mathbb{R}^{n-1})$  of bounded linear operators from  $H$  to  $\mathbb{R}^{n-1}$  and given by,

$$\|T\|_H^2 = \sum_{i=1}^{\infty} |Tu_i|^2 \text{ and } \langle T, S \rangle_H = \sum_{i=1}^{\infty} \langle Tu_i, Su_i \rangle$$

for any orthonormal basis  $(u_i)$  of  $H$ ; the value of the Hilbert-Schmidt norm and inner product are independent of the orthonormal basis used [14].

We define the space

$$D = \{(x, T) \in N \times L(\mathbb{R}^{n-1}, \mathbb{R}^n) :$$

$f \text{ is regularly differentiable at } x \text{ in direction } T\}.$

We give  $\mathbb{R}^n \times \mathbb{R}^{n-1}$  the norm  $\|(z, w)\| = |z| + |w|$  for  $z \in \mathbb{R}^n$  and  $w \in \mathbb{R}^{n-1}$ . Define a pseudonorm  $\|\cdot\|_r$  on the space  $\text{Lip}(\mathbb{R}^n \times \mathbb{R}^{n-1}, \mathbb{R}^{n-1})$  of  $\mathbb{R}^{n-1}$  valued Lipschitz functions on  $\mathbb{R}^n \times \mathbb{R}^{n-1}$  by

$$\|h\|_r = \sup_{\substack{z \in \mathbb{R}^n, u \in \mathbb{R}^{n-1} \\ |z| + |u| > 0}} \frac{|h(z, u) - h(z, 0)|}{|z| + |u|}.$$

It is easy to see this pseudonorm is lower semicontinuous in the topology of pointwise convergence and  $\|h\|_r \leq \text{Lip}(h)$ .

For  $x \in \mathbb{R}^n$  and  $T \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  we define  $f_{x,T}: \mathbb{R}^n \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$  by

$$f_{x,T}(z, u) = f(x + z + Tu).$$

The transformation  $(x, T) \mapsto f_{x,T}$  maps  $\mathbb{R}^n \times L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  to the space  $\text{Lip}(\mathbb{R}^n \times \mathbb{R}^{n-1}, \mathbb{R}^{n-1})$ . We have the inequalities

$$\|f_{x,T} - f_{y,S}\|_r \leq \text{Lip}(f)(\|T\| + \|S\|)$$

and

$$\|f_{x,T} - f_{x,S}\|_r \leq \text{Lip}(f)\|T - S\|.$$

Recall  $N$  is a  $G_\delta$  set in  $\mathbb{R}^n$ . Hence, by Lemma 5.2, there is a metric  $d_N$  on  $N$  such that  $(N, d_N)$  is complete and  $d_N$  is topologically equivalent to the Euclidean metric on  $N$ . Further, by replacing  $d_N$  by the metric  $(x, y) \mapsto d_N(x, y) + |x - y|$  if necessary, we may assume  $|x - y| \leq d_N(x, y)$  for all  $x, y \in N$ .

The space  $D$  has a standard topology as a subset of  $\mathbb{R}^n \times L(\mathbb{R}^{n-1}, \mathbb{R}^n)$ . We also define a metric, giving a different topology, on  $D$  by

$$d((x, T), (y, S)) = d_N(x, y) + \|T - S\| + \|f'(x; T) - f'(y; S)\|_H + \|f_{x,T} - f_{y,S}\|_r.$$

Note the metric  $d$  is similar to the one in [14] with the norm distance replaced by  $d_N$  and the regularity component  $\|\cdot\|_r$  slightly simplified.

We now establish several properties of  $(D, d)$ .

**Lemma 6.7.** *The following facts hold:*

(1) *The function  $(x, T) \mapsto f'(x; T)$  is  $d$ -continuous on  $D$ .*

- (2) The function  $((x, T), (y, S)) \mapsto \|f_{x,T} - f_{y,S}\|_r$  is lower semicontinuous in the standard topology of  $D \times D$ .
- (3) The function  $((x, T), (y, S)) \mapsto \|f_{x,T} - f_{y,T}\|_r$  is lower semicontinuous in the standard topology of  $D \times D$ , and so also in the topology of the product  $(D, d) \times (D, d)$ .
- (4) The space  $(D, d)$  is separable.

*Proof.* (1) follows from the definition of  $d$ .

Since  $f$  is continuous, for each fixed  $z \in \mathbb{R}^n$  and  $u \in \mathbb{R}^{n-1}$  with  $|z| + |u| > 0$  the function

$$((x, T), (y, S)) \mapsto \frac{|f_{x,T}(z, u) - f_{y,S}(z, u) - f_{x,T}(z, 0) + f_{y,S}(z, 0)|}{|z| + |u|}$$

is continuous in the standard topology of  $D \times D$ . The function

$$((x, T), (y, S)) \mapsto \|f_{x,T} - f_{y,S}\|_r$$

is a supremum of a family of continuous functions, hence lower semicontinuous. This proves (2).

The function

$$((x, T), (y, S)) \mapsto \|f_{x,T} - f_{y,T}\|_r$$

is lower semicontinuous in the standard topology of  $D \times D$  by a similar argument to that used in (2). Since the topology of  $(D, d) \times (D, d)$  has more open sets than the standard topology of  $D \times D$ , (3) follows easily.

Let  $H_0$  be the space of continuous functions  $\varphi: \mathbb{R}^n \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$ . We equip  $H_0$  with topology generated by a countable family of pseudonorms

$$\|\varphi\|_k = \sup_{\substack{z \in \mathbb{R}^n, u \in \mathbb{R}^{n-1} \\ |z| + |u| < k}} |\varphi(z, u)|.$$

The space  $H_0$  can be metrized by the metric

$$\rho_0(\varphi, \psi) = \sum_{k=1}^{\infty} 2^{-k} \min\{1, \|\varphi - \psi\|_k\}$$

and a countable dense subset of  $H_0$  is given by polynomials with rational coefficients.

Hence the space

$$H = \mathbb{R}^n \times L(\mathbb{R}^n) \times L(\mathbb{R}^{n-1}, \mathbb{R}^{n-1}) \times H_0$$

is also metrizable and separable. We show  $(D, d)$  is homeomorphic to a topological subspace of  $H$ , which implies, since subspaces of separable metric spaces are separable,  $(D, d)$  is separable.

For  $(x, T) \in D$  define  $\psi_{x,T}: \mathbb{R}^n \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$  by

$$\psi_{x,T}(z, u) = \frac{f_{x,T}(z, u) - f_{x,T}(z, 0) - f'(x; Tu)}{|z| + |u|}, \quad \psi_{x,T}(0, 0) = 0.$$

Since  $f$  is regularly differentiable at  $x$  in direction  $T$  it follows that  $\psi_{x,T}$  is continuous at  $(0, 0)$ . Hence  $\psi_{x,T} \in H_0$  so the map  $\eta: D \rightarrow H$  given by

$$\eta(x, T) = (x, T, f'(x; T), \psi_{x,T})$$

is a injection from  $D$  into  $H$ . We claim  $\eta$  is a homeomorphism onto its image. From the definition of  $d$  it is clear the first three components of  $\eta$  are  $d$  continuous. Continuity of the last component follows from the inequality

$$\|\psi_{x,T} - \psi_{y,S}\|_k \leq \|f_{x,T} - f_{y,S}\|_r + \|f'(x; T) - f'(y; S)\|.$$

Continuity of  $\eta^{-1}$  follows from the estimate

$$\begin{aligned} \|f_{x,T} - f_{y,S}\|_r &\leq \sup_{\substack{z \in \mathbb{R}^n, u \in \mathbb{R}^{n-1} \\ |z| + |u| \geq k}} \frac{|f_{x,T}(z, u) - f_{x,T}(z, 0) - f_{y,S}(z, u) + f_{y,S}(z, 0)|}{|z| + |u|} \\ &\quad + \|\psi_{x,T} - \psi_{y,S}\|_k + \|f'(x; T) - f'(y; S)\| \\ &\leq \frac{4\|f\|_\infty}{k} + \|\psi_{x,T} - \psi_{y,S}\|_k + \|f'(x; T) - f'(y; S)\| \end{aligned}$$

for every  $k \geq 1$ . This completes the proof of (4).  $\square$

**Lemma 6.8.** *The pair  $(D, d)$  is a complete metric space.*

*Proof.* Suppose  $\varepsilon_i \downarrow 0$  and  $(x_i, T_i) \in D$  are such that

$$d((x_j, T_j), (x_i, T_i)) < \varepsilon_i \text{ whenever } j \geq i.$$

Since  $(N, d_N)$  is complete it follows  $x_i$  converges to some  $x \in N$ . Clearly, from the definition of  $d$ ,  $T_i$  converges to some  $T \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  and the maps  $L_i = f'(x_i; T_i)$  converge to some  $L \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$ . Using the definition of  $d$ ,  $\|L - L_i\| \leq \varepsilon_i$ . Also, using Lemma 6.7 (2) and the definition of  $d$ ,  $\|f_{x,T} - f_{x_i,T_i}\|_r \leq \varepsilon_i$ .

To complete the proof we must show  $(x, T) \in D$  and  $(x_i, T_i)$  converges to  $(x, T)$  with respect to the metric  $d$ . To show  $(x, T) \in D$  we show  $f$  is regularly differentiable at  $x$  in direction  $T$  with  $f'(x; T) = L$ . Suppose  $\varepsilon > 0$  and find  $i$  such that  $\varepsilon_i < \varepsilon/3$ . Since  $f$  is regularly differentiable at  $x_i$  in direction  $T_i$ , we can find  $\delta > 0$  such that

$$|f(x_i + z + T_i v) - f(x_i + z) - L_i v| \leq \frac{\varepsilon}{3}(|z| + |v|)$$

whenever  $|z| + |v| < \delta$ . Hence, using the definition of  $\|\cdot\|_r$ , for  $|z| + |v| < \delta$ ,

$$\begin{aligned} &|f(x + z + T v) - f(x + z) - L v| \\ &\leq |f(x_i + z + T_i v) - f(x_i + z) - L_i v| + |L_i v - L v| \\ &\quad + \|f_{x,T} - f_{x_i,T_i}\|_r(|z| + |v|) \\ &\leq \frac{\varepsilon}{3}(|z| + |v|) + \varepsilon_i|v| + \varepsilon_i(|z| + |v|) \\ &\leq \varepsilon(|z| + |v|). \end{aligned}$$

Hence  $f$  is regularly differentiable at  $x$  in direction  $T$ .

Convergence of  $(x_i, T_i)$  to  $(x, T)$  in the metric  $d$  follows from the estimate  $\|f_{x,T} - f_{x_i,T_i}\|_r \leq \varepsilon_i$ .  $\square$

Since  $(D, d)$  is a complete separable metric space and the identity map from  $(D, d)$  to  $\mathbb{R}^n \times L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  is continuous, it follows a subset of  $D$  is Borel in  $(D, d)$  if and only if it is Borel in  $\mathbb{R}^n \times L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  [10]. In what follows it will be obvious functions we integrate are Borel in  $(D, d)$  so it follows by this fact that they are also Borel in the standard topology on  $\mathbb{R}^n \times L(\mathbb{R}^{n-1}, \mathbb{R}^n)$ .

Fix a Lipschitz function  $\Theta: L(\mathbb{R}^{n-1}, \mathbb{R}^n) \rightarrow [0, 1]$  which is differentiable everywhere and

$$\inf_{\|S\| > s} \Theta(S) > \Theta(0) = 0 \text{ for every } s > 0.$$

In the context of [14], existence of such a bump function, defined on a Banach space (and required to satisfy certain smoothness assumptions) was needed as a hypothesis on the Banach space. Existence of such a function in our case is clear - for example, identify  $L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  with  $\mathbb{R}^{n(n-1)}$  and let

$$\Theta(S) = \|S\|_E^2 / (1 + \|S\|_E^2)$$

where  $\|\cdot\|_E$  is the Euclidean norm.

We work in the subspace  $(D_0, d)$  of  $(D, d)$  where

$$D_0 = \{(x, T) \in D : d_N(x, x_0) \leq \eta_0, \|T - T_0\| \leq \eta_0\}.$$

Temporarily fix parameters  $0 < \lambda_i, \beta_i, \gamma_i, \sigma_i, s_i < \infty$  to be chosen later. We will apply the variational principle of Lemma 6.1 on the complete metric space  $(D_0, d)$  to the function  $h_0: D_0 \rightarrow \mathbb{R}$  given by

$$h_0(x, T) = -\|f'(x; T)\|_H^2,$$

with the perturbation functions  $F_i: D_0 \times D_0 \rightarrow [0, \infty)$ ,  $i \geq 0$ , defined by

$$\begin{aligned} F_i((x, T), (y, S)) = & \lambda_i d_N(x, y) + \beta_i \Theta(S - T) \\ & + \gamma_i \|f'(y; S) - f'(x; T)\|_H^2 + \sigma_i \Delta_i((x, T), (y, S)), \end{aligned}$$

where

$$\Delta_i((x, T), (y, S)) = \max\{0, \min\{1, \|f_{y,T} - f_{x,T}\|_r - s_i\}\}.$$

We now show the variational principle can be applied and prove estimates that will be useful later.

**Lemma 6.9.** *Suppose that for all  $i \geq 0$ ,*

$$0 < \lambda_i, \beta_i, \gamma_i, \sigma_i, \varepsilon_i < \infty, 0 \leq s_i < \infty, s_i \downarrow 0.$$

*Then  $F_i$  are non-negative lower semicontinuous functions on  $(D_0, d) \times (D_0, d)$  satisfying  $F_i((x, T), (x, T)) = 0$  and there are  $r_i \downarrow 0$  such that*

$$\inf\{F_i((x, T), (y, S)) : d((x, T), (y, S)) \geq r_i\} > 0.$$

*If, moreover,*

$$\|f'(x_0; T_0)\|_H^2 > \sup_{(x, T) \in D_0} \|f'(x; T)\|_H^2 - \varepsilon_0.$$

*then the function  $h_0$  and the perturbation scheme  $(F_i)$  satisfy the assumptions of the variational principle of Lemma 6.1 on the metric space  $(D_0, d)$ .*

*Proof.* Clearly  $F_i \geq 0$  and  $F_i((x, T), (x, T)) = 0$ . Lower semicontinuity of  $\Delta_i$  follows from Lemma 6.7 (3). The other terms of  $F_i$  are clearly continuous with respect to  $d$ .

Let  $r_0 = \infty$ . For  $i \geq 1$  let  $t_i = 1/2^{i+1}$ ; we show that  $r_i = s_i + (4 + \text{Lip}(f))t_i$  satisfy

$$\inf_{d((x, T), (y, S)) \geq r_i} F_i((x, T), (y, S)) \geq \min\{\lambda_i t_i, \gamma_i t_i^2, \sigma_i t_i, \inf_{\|L\| \geq t_i} \beta_i \Theta(L)\}.$$

This is obvious, from the definition of  $F_i$ , if  $d_N(x, y) \geq t_i$  or  $\|T - S\| \geq t_i$  or  $\|f'(x; T) - f'(y; S)\|_H \geq t_i$ . Suppose  $d((x, T), (y, S)) \geq r_i$  and none of the previous inequalities hold. Then

$$\|f_{x,T} - f_{y,S}\|_r \geq d((x, T), (y, S)) - 3t_i \geq s_i + (1 + \text{Lip}(f))t_i.$$

Hence

$$\begin{aligned} \|f_{x,T} - f_{y,T}\|_r &\geq \|f_{x,T} - f_{y,S}\|_r - \|f_{y,S} - f_{y,T}\|_r \\ &\geq s_i + (1 + \text{Lip}(f))t_i - \text{Lip}(f)\|S - T\| \\ &\geq s_i + t_i. \end{aligned}$$

Hence  $\|f_{x,T} - f_{y,T}\|_r \geq s_i + t_i$  and  $\sigma_i \Delta_i((x, T), (y, S)) \geq \sigma_i t_i$ .

Since  $r_0 = \infty$ , the final condition assumed is the only remaining requirement of the variational principle.  $\square$

## 7. APPLICATION OF THE VARIATIONAL PRINCIPLE

We now apply the variational principle, derive some resulting estimates and finally make exact choices of parameters.

Assume the parameters  $\lambda_i, \beta_i, \gamma_i, \sigma_i, s_i, \varepsilon_i, \varepsilon_0$  satisfy the assumptions of Lemma 6.9. Then the variational principle shows that  $(x_0, T_0)$  is the starting term of a sequence  $(x_j, T_j) \in D_0$  which  $d$ -converges to some  $(x_\infty, T_\infty) \in D_0$  and has the property that, denoting  $\varepsilon_\infty = 0$  and

$$h_i(x, T) = -\|f'(x; T)\|_H^2 + \sum_{j=0}^{i-1} F_j((x_j, T_j), (x, T)),$$

we have,

$$h_\infty(x_\infty, T_\infty) \leq \varepsilon_i + \inf_{(x, T) \in D_0} h_i(x, T)$$

for  $0 \leq i \leq \infty$ . Note that for  $i = \infty$  this says  $(x_\infty, T_\infty)$  is a minimum of  $h_\infty$  in  $D_0$ . Since  $(x_i, T_i) \in D$ , the derivatives  $L_i = f'(x_i; T_i)$  exist for all  $0 \leq i \leq \infty$ .

We guarantee  $h_\infty$  is finite by requiring

$$\sum_{i=0}^{\infty} (\lambda_i + \beta_i + \gamma_i + \sigma_i) < \infty.$$

**Lemma 7.1.** *For every  $i \geq 0$ ,*

$$d_N(x_\infty, x_i) \leq \frac{\varepsilon_i}{\lambda_i}, \quad \Theta(T_\infty - T_i) \leq \frac{\varepsilon_i}{\beta_i}, \quad \|L_\infty - L_i\|_H \leq \sqrt{\frac{\varepsilon_i}{\gamma_i}}.$$

*Proof.* It follows from the definition of  $h_i$  that

$$\begin{aligned} h_i(x_\infty, T_\infty) + \lambda_i d_N(x_\infty, x_i) + \beta_i \Theta(T_\infty - T_i) + \gamma_i \|L_\infty - L_i\|_H^2 \\ \leq h_\infty(x_\infty, T_\infty) \leq \varepsilon_i + h_i(x_\infty, T_\infty). \end{aligned}$$

Since all terms are non-negative this implies all three estimates.  $\square$



Our strategy will be to suppose  $f$  is not differentiable at  $x_\infty$  and show this would imply  $x_\infty$  is not a minimum of the function  $h_\infty$ . To do this we analyze the difference,

$$(7.1) \quad h_\infty(x_\infty, T_\infty) - h_\infty(x, T) = (\|f'(x; T)\|_H^2 - \|f'(x_\infty; T_\infty)\|_H^2) - \Phi(x) - \Psi(T) - \Upsilon(f'(x; T)) - \Delta(x)$$

where, for  $x \in N$ ,  $T \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  and  $L \in L(\mathbb{R}^{n-1}, \mathbb{R}^{n-1})$ ,

$$\begin{aligned} \Phi(x) &= \sum_{i=0}^{\infty} \lambda_i (d_N(x, x_i) - d_N(x_\infty, x_i)) \\ \Psi(T) &= \sum_{i=0}^{\infty} \beta_i (\Theta(T - T_i) - \Theta(T_\infty - T_i)) \\ \Upsilon(L) &= \sum_{i=0}^{\infty} \gamma_i (\|L - L_i\|_H^2 - \|L_\infty - L_i\|_H^2) \\ \Delta(x) &= \sum_{i=0}^{\infty} \sigma_i (\Delta_i((x_i, T_i), (x, T)) - \Delta_i((x_i, T_i), (x_\infty, T_\infty))). \end{aligned}$$

These functions are well defined and one sees from the definition of  $\Delta_i$  that  $\Delta$  does not depend on  $T$ . Also note the positive and finite function

$$\Theta_\infty(T) = \sum_{i=0}^{\infty} \beta_i \Theta(T - T_i)$$

is differentiable (by standard facts about uniformly converging series of functions).

We now give definitions of the parameters. For  $i \geq 1$ ,

$$\lambda_i = 2^{-i} \lambda_0, \quad \beta_i = 2^{-i} \beta_0, \quad \gamma_i = 2^{-i} \gamma_0, \quad s_i = 2^{-i} s_0.$$

Choose  $\varepsilon_0 > 0$  such that

$$\|f'(x_0, T_0)\|_H^2 > \sup_{(x, T) \in D_0} \|f'(x; T)\|_H^2 - \varepsilon_0.$$

Let  $\lambda_0 = 2\varepsilon_0/\eta_0$ ,  $\gamma_0 = 1/8$  and  $s_0 = 4$ . Then find  $\beta_0 > 0$  large enough that

$$\Theta(S) > \frac{\varepsilon_0}{\beta_0} \text{ whenever } \|S\| > \min \left\{ \frac{\eta_0}{2}, \frac{s_0}{8\text{Lip}(f)} \right\}.$$

For  $i \geq 1$  we choose  $\varepsilon_i > 0$  such that

$$\frac{\varepsilon_i}{\lambda_i} \leq \frac{\eta_0}{2}, \quad \frac{\varepsilon_i}{\gamma_i} \leq \frac{s_i^2}{64}$$

and

$$\Theta(S) > \frac{\varepsilon_i}{\beta_i} \text{ whenever } \|S\| > \min \left\{ \frac{\eta_0}{2}, \frac{s_i}{8\text{Lip}(f)} \right\}.$$

We now deduce estimates about the speed of convergence of  $x_i$ ,  $T_i$  and  $L_i$ .

**Lemma 7.2.** *For all  $0 \leq i < \infty$  we have the following estimates of the speed of convergence:*

$$d_N(x_i, x_\infty) \leq \frac{\eta_0}{2}, \quad \|T_i - T_\infty\| \leq \min \left\{ \frac{\eta_0}{2}, \frac{s_i}{8\text{Lip}(f)} \right\},$$

and

$$\|L_i - L_\infty\| \leq \frac{s_i}{8}.$$

Also, for  $0 \leq i \leq \infty$ ,

$$\|T_i\| \leq 1, \quad \|L_i\| \leq \frac{5}{4}.$$

*Proof.* The first two inequalities, for  $0 \leq i < \infty$ , and the third inequality, for  $0 < i < \infty$ , follow from the definitions and Lemma 7.1. Using the definition of  $f'(x; T)$  and the second inequality for  $i = 0$  we obtain

$$\|L_0 - L_\infty\| \leq \text{Lip}(f)\|T_0 - T_\infty\| \leq \frac{s_0}{8}$$

which is the third inequality for  $i = 0$ . Again, using the second inequality we have, for any  $0 \leq i \leq \infty$ ,

$$\|T_i\| \leq \|T_0\| + \|T_0 - T_\infty\| + \|T_i - T_\infty\| \leq \frac{1}{2} + \eta_0 \leq 1.$$

Since  $(x_i, T_i) \in D_0$  for all  $0 \leq i \leq \infty$ , the final estimate follows from the assumptions on  $f$  (assumptions 6.2).  $\square$

Finally for  $i \geq 0$  we choose  $\sigma_i > 0$  such that

$$\sigma_i \frac{C(\mu)(5\text{Lip}(f))^{2(n+C(\mu))-2}}{(s_i/8)^{2(n+C(\mu))-2}} \leq 2^{-i-3},$$

$$\sigma_i \leq 2^{-i-4},$$

and

$$\sigma_i \leq \frac{2^{-i-5}}{i+1} \frac{s_i}{48\text{Lip}(f)}.$$

## 8. NON-DIFFERENTIABILITY CONTRADICTS MINIMALITY

In this section we suppose  $f$  is not differentiable at  $x_\infty$  and deduce consequences. We then define a region in which we plan to find a point  $x$ , with corresponding direction  $T$ , such that  $(x, T) \in D_0$  and  $h_\infty(x, T) < h_\infty(x_\infty, T_\infty)$ . This contradiction would then show  $f$  must be differentiable at  $x_\infty$ . Most of the arguments are similar to those in [14] except we must choose the region mentioned so that it contains many points of  $N$  which are regular points of the function  $f$ .

Let  $L_\Theta$  be the derivative of  $\Theta_\infty$  at  $T_\infty$ . We recall

$$\begin{aligned} h_\infty(x, T) = & -\|f'(x; T)\|_H^2 + \sum_{i=0}^{\infty} \lambda_i d_N(x, x_i) + \sum_{i=0}^{\infty} \beta_i \Theta(T - T_i) \\ & + \sum_{i=0}^{\infty} \gamma_i \|f'(x; T) - L_i\|_H^2 \\ & + \sum_{i=0}^{\infty} \sigma_i \max\{0, \min\{1, \|f_{x, T_i} - f_{x_i, T_i}\|_r\} - s_i\}. \end{aligned}$$

Suppose for the moment  $f$  is differentiable at  $x_\infty$ . Since  $h_\infty$  has a minimum at  $(x_\infty, T_\infty)$ , this implies that the function  $T \mapsto h_\infty(x_\infty, T)$  is differentiable with derivative zero at  $T_\infty$ . We differentiate to find equations for  $f'(x_\infty; S)$ . Observe

$$\|A_0 + A\|_H^2 - \|A_0\|_H^2 - 2\langle A_0, A \rangle_H = \|A\|_H^2$$

for linear maps  $A, A_0 \in L(\mathbb{R}^{n-1}, \mathbb{R}^{n-1})$ . This implies the derivative of the map

$$A \mapsto \|A\|_H^2$$

at  $A_0$  is the map

$$A \mapsto 2\langle A_0, A \rangle_H.$$

Hence, using the chain rule and the fact  $T \mapsto f'(x; T)$  is linear, we deduce

$$(8.1) \quad -2\langle L_\infty, f'(x_\infty; S) \rangle_H + L_\Theta S + 2 \sum_{i=0}^{\infty} \gamma_i \langle L_\infty - L_i, f'(x_\infty; S) \rangle_H = 0$$

for all  $S \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$ . We now solve this system of equations for an unknown operator  $L \in L(\mathbb{R}^n, \mathbb{R}^{n-1})$  in place of  $f'(x_\infty)$ . Denote

$$R_\infty = 2L_\infty + 2 \sum_{i=0}^{\infty} \gamma_i (L_i - L_\infty) \in L(\mathbb{R}^{n-1}, \mathbb{R}^{n-1}).$$

Then we can rewrite the system of equations as

$$(8.2) \quad \langle R_\infty, LS \rangle_H = L_\Theta S \text{ for all } S \in L(\mathbb{R}^{n-1}, \mathbb{R}^n).$$

Since every  $S \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  is a sum of rank one operators, we may consider this linear equation only for rank one operators  $S \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$ . We first deduce  $R_\infty$  is invertible.

**Lemma 8.1.** *The operator  $R_\infty$  is invertible and  $\|R_\infty\|^{-1} \leq 1$ .*

*Proof.* By Lemma 7.2 we know  $\|L_i - L_\infty\| \leq 1/2$  for all  $i \geq 0$ . Since  $\sum_{i=0}^{\infty} \gamma_i = 1/4$ , and  $\|\text{Id} - L_\infty\| \leq 1/4$  by assumptions 6.2, we obtain,

$$\begin{aligned} \|\text{Id} - (1/2)R_\infty\| &= \|\text{Id} - L_\infty - \sum_{i=0}^{\infty} \gamma_i (L_i - L_\infty)\| \\ &\leq \|\text{Id} - L_\infty\| + \sum_{i=0}^{\infty} \gamma_i \|L_i - L_\infty\| \\ &\leq 1/2. \end{aligned}$$

This implies the expression

$$\sum_{i=0}^{\infty} (\text{Id} - (1/2)R_{\infty})^i$$

is a well defined linear map with norm at most 2. An easy computation then shows it is the inverse of  $(1/2)R_{\infty}$ . Hence the map  $R_{\infty}$  is invertible with  $\|R_{\infty}^{-1}\| \leq 1$ .  $\square$

A rank one operator  $S \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  can be written as  $S = x \otimes e$  where  $x \in \mathbb{R}^n$  and  $e \in \mathbb{R}^{n-1}$ . It acts by  $(x \otimes e)(u) = \langle e, u \rangle x$ . We note  $L(x \otimes e) = Lx \otimes e$  for a linear map  $L$  defined on  $\mathbb{R}^n$ . Indeed, for any  $u \in \mathbb{R}^{n-1}$ ,

$$L(x \otimes e)u = L(\langle e, u \rangle x) = \langle e, u \rangle Lx = (Lx \otimes e)(u).$$

Also  $\langle R, w \otimes e \rangle_H = \langle Re, w \rangle$  for any  $w \in \mathbb{R}^n$  and  $R \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$ . To see this let  $e_i$  be an orthonormal basis of  $\mathbb{R}^{n-1}$ . Then,

$$\langle R, w \otimes e \rangle_H = \sum_{i=1}^{n-1} \langle R(e_i), \langle e_i, e \rangle w \rangle = \langle R(\sum_{i=1}^{n-1} \langle e_i, e \rangle e_i), w \rangle = \langle Re, w \rangle.$$

Using these facts, the system of equations (8.2) for the unknown  $Lx$  can be written

$$\langle R_{\infty}e, Lx \rangle = L_{\Theta}(x \otimes e), \text{ for all } e \in \mathbb{R}^{n-1}.$$

Since  $R_{\infty}$  is invertible this can be written as

$$(8.3) \quad \langle e, Lx \rangle = L_{\Theta}(x \otimes (R_{\infty}^{-1}e)) \text{ for all } e \in \mathbb{R}^{n-1}.$$

For a fixed  $x$ , the map  $e \mapsto L_{\Theta}(x \otimes (R_{\infty}^{-1}e))$  defines a linear functional on  $\mathbb{R}^{n-1}$  and the left hand side of (8.3) says this linear functional is represented by  $Lx \in \mathbb{R}^{n-1}$ . Hence  $Lx$  is uniquely defined and linearity of (8.3) in  $x$  shows that  $L$  is linear. Further,  $L$  is bounded as  $|Lx| \leq \|L_{\Theta}\| \|R_{\infty}^{-1}\| |x|$ . Hence we have shown the following lemma.

**Lemma 8.2.** *The system of equations (8.2) uniquely defines an operator  $L \in L(\mathbb{R}^n, \mathbb{R}^{n-1})$  with  $\|L\| \leq \|L_{\Theta}\| \|R_{\infty}^{-1}\| \leq \|L_{\Theta}\|$ .*

By revisiting (8.1) we see that, even if  $f'(x_{\infty})$  does not exist, the argument above still shows the following fact.

**Lemma 8.3.**  $f'(x_{\infty}, T_{\infty}) = LT_{\infty}$ .

In order to prove Proposition 6.3 it suffices to prove the following proposition.

**Proposition 8.4.**  *$f$  is differentiable at  $x_{\infty} \in N$  and  $f'(x_{\infty}) = L$ .*

To prove Proposition 8.4 we argue by contradiction. Assume that  $L$  is not the derivative of  $f$  at  $x_{\infty}$ .

**Lemma 8.5.** *There is  $\varepsilon > 0$  such that for every  $\delta > 0$  one may find  $x \in \mathbb{R}^n$  with  $|x| < \delta$  such that*

- $Lx = 0$ .
- $|f(x_{\infty} + x) - f(x_{\infty})| > \varepsilon|x|$ .

*Proof.* Let  $c = 1 + \|T_\infty L_\infty^{-1} L\| + \|L_\infty^{-1} L\|$ . Since  $L$  is not the derivative of  $f$  at  $x_\infty$ , there is  $\varepsilon > 0$  such that for every  $\delta > 0$  there exists  $\tilde{x} \in \mathbb{R}^n$  with  $|\tilde{x}| < \delta$  and

$$(8.4) \quad |f(x_\infty + \tilde{x}) - f(x_\infty) - L\tilde{x}| > 2\varepsilon c|\tilde{x}|.$$

Suppose  $\delta > 0$  is fixed. Since  $f$  is regularly differentiable at  $x_\infty$  in direction  $T_\infty$  there exists  $\eta > 0$  such that

$$(8.5) \quad |f(x_\infty + z + T_\infty u) - f(x_\infty + z) - L_\infty u| \leq \varepsilon(|z| + |u|)$$

whenever  $z \in \mathbb{R}^n, u \in \mathbb{R}^{n-1}$  and  $|z| + |u| < \eta$ .

Choose  $\tilde{x} \in \mathbb{R}^n$  such that  $c|\tilde{x}| < \min\{\delta, \eta\}$  and (8.4) holds. Define the point  $x = \tilde{x} - T_\infty L_\infty^{-1} L\tilde{x}$ . Since  $LT_\infty = L_\infty$  by Lemma 8.3, it follows  $Lx = 0$ . Further,

$$|x| + |L_\infty^{-1} L\tilde{x}| \leq |\tilde{x}| + |T_\infty L_\infty^{-1} L\tilde{x}| + |L_\infty^{-1} L\tilde{x}| \leq c|\tilde{x}| < \min\{\delta, \eta\}.$$

Hence we may apply (8.5) with  $z = x$  and  $u = L_\infty^{-1} L\tilde{x}$  to obtain

$$\begin{aligned} & |f(x_\infty + \tilde{x}) - f(x_\infty + x) - L\tilde{x}| \\ &= |f(x_\infty + x + T_\infty L_\infty^{-1} L\tilde{x}) - f(x_\infty + x) - L\tilde{x}| \\ &\leq \varepsilon(|x| + |L_\infty^{-1} L\tilde{x}|) \leq \varepsilon c|\tilde{x}|. \end{aligned}$$

By combining this with (8.4) and using the triangle inequality we obtain

$$\begin{aligned} & |f(x_\infty + x) - f(x_\infty)| \\ &\geq |f(x_\infty + \tilde{x}) - f(x_\infty) - L\tilde{x}| - |f(x_\infty + \tilde{x}) - f(x_\infty + x) - L\tilde{x}| \\ &> \varepsilon c|\tilde{x}| \geq \varepsilon|x| \end{aligned}$$

as required.  $\square$

We now define various parameters and a region in which to look for a pair  $(x, T) \in D_0$  with  $h_\infty(x, T) < h_\infty(x_\infty, T_\infty)$ . We fix  $0 < \varepsilon < 1$  with the property from Lemma 8.5. By differentiability of  $\Theta_\infty$  we find  $K \geq 1/\varepsilon$  such that the parameter

$$\kappa = \frac{s_K}{48\text{Lip}(f)}$$

satisfies  $\kappa < \eta_0/4$  and

$$(8.6) \quad \Theta_\infty(T_\infty + S) - \Theta_\infty(T_\infty) \leq L_\Theta S + \frac{\varepsilon}{8}\|S\|$$

whenever  $\|S\| \leq \kappa$ . Let  $\tau = \varepsilon\kappa/4$ ,

$$t = \frac{\tau}{22n\text{Lip}(f)(1 + \|R_\infty\|_H)},$$

and, using Lemma 2.3, choose  $0 < \xi \leq t/6$  such that for every  $\theta > 0$  we have

$$\frac{\mu(B(0, \theta(1 - 24\xi/ts_K)))}{\mu(B(0, \theta))} (1 - 24n\xi/s_K) \geq 1 - \tau.$$

Since  $f$  is regularly differentiable at each  $x_i$  in the direction of  $T_i$ , there is  $\delta_1 > 0$  such that for all  $i = 0, \dots, K$  and  $i = \infty$ ,

$$(8.7) \quad |f(x_i + z + T_i v) - f(x_i + z) - L_i v| \leq \xi(|z| + |v|)$$

whenever  $z \in \mathbb{R}^n, v \in \mathbb{R}^{n-1}$  and  $|z| + |v| < \delta_1$ .

Let  $\delta_2 > 0$  be such that

$$\delta_2 \leq \min \left\{ \frac{\kappa\delta_1}{6}, \frac{\kappa\eta_0}{6}, \frac{\delta_1\kappa\tau}{24\text{Lip}(f)} \right\}$$

and  $z \in N$ ,  $|z - x_\infty| \leq 3\delta_2/\kappa$  implies

$$d_N(z, x_\infty) \leq \min \left\{ \frac{\tau}{\sum_{i=0}^\infty \lambda_i}, \frac{\eta_0}{2} \right\}.$$

Use Lemma 8.5 to find  $x \in \mathbb{R}^n$  such that  $|x| < \delta_2$ ,  $Lx = 0$  and

$$|f(x_\infty + x) - f(x_\infty)| > \varepsilon|x|.$$

Denote

$$r = \frac{|x|}{\kappa}, \quad w = \frac{1}{r}(f(x_\infty + x) - f(x_\infty)), \quad e = \frac{R_\infty^{-1}w}{|R_\infty^{-1}w|},$$

and

$$\Omega = \{u \in \mathbb{R}^{n-1} : -r < \langle e, u \rangle < 0, |u - \langle e, u \rangle e| < r\}.$$

We have, since  $\|R_\infty^{-1}\| \leq 1$ ,

$$\langle R_\infty e, w \rangle = \frac{|w|^2}{|R_\infty^{-1}w|} \geq |w| > \varepsilon\kappa = 4\tau.$$

Temporarily fix  $\tilde{e} \in \mathbb{R}^{n-1}$ ,  $\tilde{x} \in \mathbb{R}^n$  and  $\tilde{T}_\infty \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  close to  $e$ ,  $x$  and  $T_\infty$  respectively. Define an affine map  $\gamma: \mathbb{R}^{n-1} \rightarrow \mathbb{R}^n$  with  $\text{Lip}(\gamma) \leq 2\kappa$  by

$$\gamma(u) = \tilde{x} + \langle \tilde{e}, u \rangle \tilde{x}/r$$

and let  $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$  be an affine change of coordinates given by

$$\varphi(u) = x_\infty + \gamma(\pi_{n-1}u) + \tilde{T}_\infty \pi_{n-1}u + (0, \pi^{n-1}u)$$

where  $\pi_{n-1}$  and  $\pi^{n-1}$  are the orthogonal projections onto the first  $n$  coordinates and the final coordinate respectively. That is, for  $a \in \mathbb{R}^{n-1}$  and  $b \in \mathbb{R}$ , we have  $\pi_{n-1}(a, b) = a$  and  $\pi^{n-1}(a, b) = b$ .

Note that,

$$\varphi(u) = x_\infty + \tilde{x} + \kappa \langle \tilde{e}, u \rangle \frac{\tilde{x}}{|x|} + \tilde{T}_\infty \pi_{n-1}u + (0, \pi^{n-1}u)$$

and

$$\varphi'_{n-1}(u) = \kappa \frac{\tilde{x} \otimes \tilde{e}}{|x|} + \tilde{T}_\infty$$

for  $u \in \mathbb{R}^n$ . Let  $Q = \Omega \times (-tr, tr)$ .

**Lemma 8.6.** *Let  $\alpha > 0$ . Then there exists  $\tilde{e} \in \mathbb{R}^{n-1}$ ,  $\tilde{x} \in \mathbb{R}^n$ , and a linear map  $\tilde{T}_\infty \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  such that*

$$|\tilde{e} - e| + |\tilde{x} - x| + \|\tilde{T}_\infty - T_\infty\| < \alpha$$

and, for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ ,

- The function  $f$  is regularly differentiable at  $\varphi(u)$  in the direction  $\varphi'_{n-1}(u)$ ,
- $\varphi(u) \in N$ .

*Proof.* We can choose  $\tilde{e}, \tilde{x}$  and  $\tilde{T}_\infty$  with

$$|\tilde{e} - e| + |\tilde{x} - x| + \|\tilde{T}_\infty - T_\infty\| < \alpha$$

so that  $\varphi$  is bilipschitz and belongs to the countable dense family of affine maps in Definition 5.3. This implies  $\varphi(u) \in N$  whenever  $u_n \in \tilde{N}$ . Hence, since  $\mu(\mathbb{R} \setminus \tilde{N}) = 0$ , it follows  $\varphi(u) \in N$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ .

The preimage, under the bilipschitz map  $\varphi$ , of the  $\sigma$ -porous set of points at which  $f$  is not regular is again  $\sigma$ -porous. Hence, since  $\mathcal{L}^{n-1} \times \mu$  is doubling, Lemma 4.5 and Lemma 4.6 imply that the function  $f$  is regular at  $\varphi(u)$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u$ .

By the classical Rademacher theorem we know that for each fixed  $b \in \mathbb{R}$ ,  $f$  is differentiable at  $\varphi(a, b)$  in direction  $\varphi'_{n-1}(a, b)$  for  $\mathcal{L}^{n-1}$  almost every  $a \in \mathbb{R}^{n-1}$ . Hence  $f$  is differentiable at  $\varphi(u)$  in direction  $\varphi'_{n-1}(u)$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ . By Lemma 4.3 this proves the result.  $\square$

We now fix  $\alpha > 0$  small relative to all previous parameters. Since  $\alpha$  is the last parameter we define, we don't list precise estimates. When the fact  $\alpha$  is small is used, it should be clear  $\alpha$  could have been chosen appropriately at this stage. Fix  $\tilde{e} \in \mathbb{R}^{n-1}$ ,  $\tilde{x} \in \mathbb{R}^n$  and  $\tilde{T}_\infty \in L(\mathbb{R}^{n-1}, \mathbb{R}^n)$  as in Lemma 8.6 with corresponding  $\gamma$  and  $\varphi$ . Define  $g: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  by

$$g(u) = f(\varphi(u)) - L_\infty \pi_{n-1} u.$$

**Lemma 8.7.** *If  $\alpha$  is sufficiently small then  $\text{Lip}(\varphi) \leq 3$ ,  $\text{Lip}(g) \leq 5\text{Lip}(f)$ ,*

- $|\varphi(u) - x_\infty| \leq 3r$  for all  $u \in Q$ , and
- for every  $u, v \in \overline{Q}$  with  $\langle u - v, e \rangle = 0$ ,

$$|g(u) - g(v)| \leq \frac{\tau r}{n(1 + \|R_\infty\|_H)}.$$

*Proof.* We use Lemma 7.2. For the first inequality, since  $\|T_\infty\| \leq 1$  and  $\alpha$  is small,

$$\text{Lip}(\varphi) \leq 2\kappa + \|\tilde{T}_\infty\| + 1 \leq 3.$$

For the second inequality, the facts  $\|L_\infty\| \leq 5/4$  and  $\text{Lip}(f) \geq 1$  immediately imply

$$\text{Lip}(g) \leq \frac{5}{4} + 3\text{Lip}(f) \leq 5\text{Lip}(f).$$

For the third inequality, for small  $\alpha$ , we estimate,

$$|\varphi(u) - x_\infty| \leq \kappa r + 2\kappa r + 2r + tr + \kappa r = (4\kappa + t + 2)r \leq 3r.$$

For the final estimate we use regular differentiability of  $f$  at  $x_\infty$  in direction  $T_\infty$  which is (8.7). For convenience, first assume  $\tilde{e} = e, \tilde{x} = x$  and  $\tilde{T}_\infty = T_\infty$ .

We next suppose  $u, v \in \overline{\Omega} \subset \mathbb{R}^{n-1} \subset \mathbb{R}^n$ . Since  $\langle u - v, e \rangle = 0$  it follows that  $\gamma(\pi_{n-1}u) = \gamma(\pi_{n-1}v)$  and hence  $\varphi(u) - \varphi(v) = T_\infty(u - v)$ . We use (8.7) with  $z = \gamma(v) + T_\infty v$  and  $u - v$  in place of  $u$ . We estimate,

$$|\gamma(v) + T_\infty v| + |u - v| \leq |\varphi(v) - x_\infty| + |u - v| \leq 6r < \delta_1.$$

Hence, using (8.7) and recalling  $\gamma(u) = \gamma(v)$ ,

$$\begin{aligned}
& |g(u) - g(v)| \\
&= |f(x_\infty + \gamma(u) + T_\infty u) - f(x_\infty + \gamma(v) + T_\infty v) - L_\infty(u - v)| \\
&= |f(x_\infty + \gamma(v) + T_\infty v + T_\infty(u - v)) - f(x_\infty + \gamma(v) + T_\infty v) - L_\infty(u - v)| \\
&\leq \xi(|\gamma(v) + T_\infty v| + |u - v|) \\
&\leq 6\xi r.
\end{aligned}$$

For general  $u, v \in \overline{Q}$  we just use that  $Q$  is relatively thin in the remaining direction,

$$\begin{aligned}
|g(u) - g(v)| &\leq |g(u) - g(\pi_{n-1}u)| + |g(v) - g(\pi_{n-1}v)| \\
&\quad + |g(\pi_{n-1}u) - g(\pi_{n-1}v)| \\
&\leq 2\text{Lip}(g)tr + 6\xi r \leq 10\text{Lip}(f)tr + tr \\
&\leq 11\text{Lip}(f)tr = \frac{\tau r}{2n(1 + \|R_\infty\|_H)}.
\end{aligned}$$

For general  $\tilde{e}, \tilde{x}$  and  $\tilde{T}_\infty$  we note that, if we temporarily denote  $g = g_{\tilde{e}, \tilde{x}, \tilde{T}_\infty}$  then since  $f$  is Lipschitz, provided  $\alpha$  is sufficiently small, we can ensure

$$|g_{\tilde{e}, \tilde{x}, \tilde{T}_\infty}(u) - g_{e, x, T_\infty}(u)| \leq \frac{\tau r}{4n(1 + \|R_\infty\|_H)}$$

for all  $u \in \overline{Q}$ . Thus, provided  $\alpha$  is sufficiently small, the result follows.  $\square$

We now observe that  $(\varphi(u), \varphi'_{n-1}(u)) \in D_0$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ . Indeed, for all  $u \in Q$  we have  $|\varphi(u) - x_\infty| \leq 3r$ . If  $\varphi(u) \in N$  this implies  $d_N(\varphi(u), x_\infty) \leq \eta_0/2$ . Hence, since  $d_N(x_\infty, x_0) < \eta_0/2$  by Lemma 7.2,

$$d_N(\varphi(u), x_0) \leq d_N(\varphi(u), x_\infty) + d_N(x_\infty, x_0) \leq \eta_0.$$

Further, using the expression for  $\varphi'_{n-1}$  and Lemma 7.2, it follows that for sufficiently small  $\alpha$ ,

$$\|\varphi'_{n-1}(u) - T_0\| \leq \|\varphi'_{n-1}(u) - \tilde{T}_\infty\| + \|\tilde{T}_\infty - T_\infty\| + \|T_\infty - T_0\| \leq \eta_0.$$

By our choice of  $\varphi$  after Lemma 8.6,  $\varphi(u) \in N$  and  $f$  is regularly differentiable at  $\varphi(u)$  in direction  $\varphi'_{n-1}(u)$  at  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ . Hence  $(\varphi(u), \varphi'_{n-1}(u)) \in D_0$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ .

We aim to show that

$$(8.8) \quad \int_Q (h_\infty(x_\infty, T_\infty) - h_\infty(\varphi(u), \varphi'_{n-1}(u))) \, d(\mathcal{L}^{n-1} \times \mu) > 0.$$

Once this is done a contradiction follows. Indeed, suppose (8.8) holds. Then there is  $u \in Q$  such that  $(\varphi(u), \varphi'_{n-1}(u)) \in D_0$  and

$$h_\infty(x_\infty, T_\infty) - h_\infty(\varphi(u), \varphi'_{n-1}(u)) > 0$$

which contradicts the assumption  $(x_\infty, T_\infty)$  is a minimizer of  $h_\infty$  in  $D_0$ .



## 9. INTEGRAL ESTIMATES

We now prove (8.8). Recall from (7.1),

$$\begin{aligned}
 (9.1) \quad & h_\infty(x_\infty, T_\infty) - h_\infty(\varphi, \varphi'_{n-1}) \\
 &= (\|f'(\varphi; \varphi'_{n-1})\|_H^2 - \|L_\infty\|_H^2) \\
 &\quad - \Phi(\varphi) - \Psi(\varphi'_{n-1}) - \Upsilon(f'(\varphi; \varphi'_{n-1})) - \Delta(\varphi).
 \end{aligned}$$

We now estimate the average integral of each of the terms on the right side. Most of the estimates are very similar to those in [14]. The main difference is in the estimate of the regularity term. Here we need to account for the fact  $\mu$  is not Lebesgue measure but can still make the necessary estimates using Proposition 3.3.

**9.1. Estimate of  $\int_Q \Phi(\varphi) \, d(\mathcal{L}^{n-1} \times \mu)$ .** For  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$  we know, by Lemma 8.6,  $\varphi(u) \in N$ . By Lemma 8.7 we know  $|\varphi(u) - x_\infty| \leq 3r$  for every  $u \in \overline{Q}$ . By our choice of  $r$  this implies  $d_N(\varphi(u), x_\infty) \leq \tau / (\sum_{i=0}^\infty \lambda_i)$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u \in Q$ .

Hence, for such  $u$ ,

$$|d_N(x_i, \varphi(u)) - d_N(x_i, x_\infty)| \leq d_N(\varphi(u), x_\infty) \leq \tau / \sum_{i=0}^\infty \lambda_i.$$

This implies,

$$|\Phi(\varphi(u))| \leq \sum_{i=0}^\infty \lambda_i d_N(\varphi(u), x_\infty) \leq \tau.$$

Hence

$$\int_Q \Phi(\varphi) \, d(\mathcal{L}^{n-1} \times \mu) \leq \tau.$$

**9.2. Estimate of  $\int_Q \Psi(\varphi'_{n-1}) \, d(\mathcal{L}^{n-1} \times \mu)$ .** By our choice of  $x$  we know  $Lx = 0$ . By the definition of  $L$  this implies  $L_\Theta(x \otimes e) = 0$ . Using (8.6) we obtain, provided  $\alpha$  is sufficiently small,

$$\begin{aligned}
 \Psi(\varphi'_{n-1}) &= \Theta_\infty \left( \tilde{T}_\infty + \kappa \frac{\tilde{x} \otimes \tilde{e}}{|\tilde{x}|} \right) - \Theta_\infty(T_\infty) \\
 &\leq \Theta_\infty \left( T_\infty + \kappa \frac{x \otimes e}{|x|} \right) - \Theta_\infty(T_\infty) \\
 &\quad + \text{Lip}(\Theta_\infty) \left\| \tilde{T}_\infty + \kappa \frac{\tilde{x} \otimes \tilde{e}}{|\tilde{x}|} - T_\infty - \kappa \frac{x \otimes e}{|x|} \right\| \\
 &\leq \frac{\varepsilon \kappa}{8} + \frac{\varepsilon \kappa}{8} \\
 &= \tau.
 \end{aligned}$$

Hence

$$\int_Q \Psi(\varphi'_{n-1}) \, d(\mathcal{L}^{n-1} \times \mu) \leq \tau.$$

**9.3. Estimate of  $\int_Q \Upsilon(f'(\varphi; \varphi'_{n-1})) \, d(\mathcal{L}^{n-1} \times \mu)$ .** Using the definition of  $g$  we have  $f'(\varphi; \varphi'_{n-1}) = g'_{n-1} + L_\infty$ . Since  $\sum_{i=0}^\infty \gamma_i = 1/4$  we have,

$$\begin{aligned} \Upsilon(f'(\varphi(u); \varphi'_{n-1}(u))) &= \sum_{i=0}^\infty \gamma_i (\| (L_i - L_\infty) - g'_{n-1}(u) \|_H^2 - \| L_i - L_\infty \|_H^2) \\ &= \sum_{i=0}^\infty \gamma_i (2 \langle (L_\infty - L_i), g'_{n-1}(u) \rangle_H + \| g'_{n-1}(u) \|_H^2) \\ &= \left\langle 2 \sum_{i=0}^\infty \gamma_i (L_\infty - L_i), g'_{n-1}(u) \right\rangle_H + \frac{1}{4} \| g'_{n-1}(u) \|_H^2. \end{aligned}$$

Hence

$$\begin{aligned} \int_Q \Upsilon(f'(\varphi; \varphi'_{n-1})) \, d(\mathcal{L}^{n-1} \times \mu) &= \int_Q \left\langle 2 \sum_{i=0}^\infty \gamma_i (L_\infty - L_i), g'_{n-1}(u) \right\rangle_H \, d(\mathcal{L}^{n-1} \times \mu) \\ &\quad + \frac{1}{4} \int_Q \| g'_{n-1} \|_H^2 \, d(\mathcal{L}^{n-1} \times \mu). \end{aligned}$$

**9.4. Estimate of  $\int_Q \Delta(\varphi) \, d(\mathcal{L}^{n-1} \times \mu)$ .** First fix  $0 \leq i \leq K$ . Denote  $s = 24\xi r/s_K$  and

$$P = \{u \in Q : \Delta_i((x_i, T_i), (\varphi(u), \varphi'_{n-1}(u))) - \Delta_i((x_i, T_i), (x_\infty, T_\infty)) > \tau\}.$$

We show that for every  $u \in P$ ,

$$(9.2) \quad \text{reg}_{n-1} g(u, B(u, s)) > \frac{s_i}{8}.$$

Fix  $u \in P$ . Choose  $\tilde{z} \in \mathbb{R}^n$  and  $\tilde{w} \in \mathbb{R}^{n-1} \subset \mathbb{R}^n$  such that

$$\begin{aligned} &|f_{x_i, T_i}(\tilde{z}, \tilde{w}) - f_{x_i, T_i}(\tilde{z}, 0) - f_{\varphi(u), T_i}(\tilde{z}, \tilde{w}) + f_{\varphi(u), T_i}(\tilde{z}, 0)| \\ &> \left( \|f_{x_i, T_i} - f_{\varphi(u), T_i}\|_r - \frac{\tau}{2} \right) (|\tilde{z}| + |\tilde{w}|). \end{aligned}$$

We show that if  $\tilde{u} = \tilde{z} + \tilde{w}$  then

$$(9.3) \quad |\tilde{z}| + |\tilde{w}| < \delta_1 \text{ and } u + \tilde{u}, u + \tilde{z} \in B(u, s).$$

We have two cases

*Case 1.* Suppose that  $|\tilde{z}| + |\tilde{w}| \geq \delta_1/2$ . From Lemma 8.7 and the choice of  $\delta_2$ ,

$$2\text{Lip}(f)|\varphi(u) - x_\infty| \leq 6\text{Lip}(f)r \leq 6\text{Lip}(f)\frac{\delta_2}{\kappa} \leq \frac{\tau}{4}\delta_1 \leq \frac{\tau}{2}(|\tilde{z}| + |\tilde{w}|).$$

It follows that

$$\begin{aligned}
& (|\tilde{z}| + |\tilde{w}|) \|f_{x_i, T_i} - f_{x_\infty, T_i}\|_r \\
& \geq |f_{x_i, T_i}(\tilde{z}, \tilde{w}) - f_{x_i, T_i}(\tilde{z}, 0) - f_{x_\infty, T_i}(\tilde{z}, \tilde{w}) + f_{x_\infty, T_i}(\tilde{z}, 0)| \\
& \geq |f_{x_i, T_i}(\tilde{z}, \tilde{w}) - f_{x_i, T_i}(\tilde{z}, 0) - f_{\varphi(u), T_i}(\tilde{z}, \tilde{w}) + f_{\varphi(u), T_i}(\tilde{z}, 0)| \\
& \quad - 2\text{Lip}(f)|\varphi(u) - x_\infty| \\
& \geq \left( \|f_{x_i, T_i} - f_{\varphi(u), T_i}\|_r - \frac{\tau}{2} \right) (|\tilde{z}| + |\tilde{w}|) - \frac{\tau}{2} (|\tilde{z}| + |\tilde{w}|) \\
& \geq (|\tilde{z}| + |\tilde{w}|) (\|f_{x_i, T_i} - f_{\varphi(u), T_i}\|_r - \tau).
\end{aligned}$$

Using the definition of  $\Delta_i$  this implies

$$\Delta_i((x_i, T_i), (\varphi(u), \varphi'_{n-1}(u))) - \Delta((x_i, T_i), (x_\infty, T_\infty)) \leq \tau$$

which contradicts the assumption that  $u \in P$ .

*Case 2.* Suppose that  $|\tilde{z}| + |\tilde{w}| < \delta_1/2$ . If  $s \geq |\tilde{u}| + |\tilde{z}|$  then (9.3) is clear. Hence we suppose  $s \leq |\tilde{u}| + |\tilde{z}|$ . It follows that  $s \leq 2(|\tilde{z}| + |\tilde{w}|)$ . Define  $\hat{z} \in \mathbb{R}^n$  by the requirement that  $\varphi(u) + \hat{z} = x_\infty + \hat{z}$ . Then by Lemma 8.7 and the choice of  $\delta_2$ ,

$$|\hat{z}| + |\tilde{w}| \leq |\varphi(u) - x_\infty| + |\tilde{z}| + |\tilde{w}| \leq 3r + |\tilde{z}| + |\tilde{w}| < \delta_1.$$

We use the estimate (8.7) from regular differentiability, our choice of  $\xi$  and  $s$ , and the fact  $\|T_i - T_\infty\| \leq s_i/8\text{Lip}(f)$  and  $\|L_i - L_\infty\| \leq s_i/8$  from Lemma 7.2 to conclude,

$$\begin{aligned}
& |f_{\varphi(u), T_i}(\tilde{z}, \tilde{w}) - f_{\varphi(u), T_i}(\tilde{z}, 0) - L_i \tilde{w}| \\
& = |f_{x_\infty, T_i}(\hat{z}, \tilde{w}) - f_{x_\infty, T_i}(\hat{z}, 0) - L_i \tilde{w}| \\
& \leq |f_{x_\infty, T_\infty}(\hat{z}, \tilde{w}) - f_{x_\infty, T_\infty}(\hat{z}, 0) - L_\infty \tilde{w}| \\
& \quad + |f_{x_\infty, T_\infty}(\hat{z}, \tilde{w}) - f_{x_\infty, T_i}(\hat{z}, \tilde{w})| + |L_i \tilde{w} - L_\infty \tilde{w}| \\
& \leq \xi(|\hat{z}| + |\tilde{w}|) + \text{Lip}(f)|T_\infty \tilde{w} - T_i \tilde{w}| + \frac{s_i}{8} |\tilde{w}| \\
& \leq \xi(3r + |\tilde{z}| + |\tilde{w}|) + \frac{s_i}{4} |\tilde{w}| \\
& \leq \frac{ss_i}{8} + \frac{s_i}{2} (|\tilde{z}| + |\tilde{w}|) \\
& \leq \frac{3s_i}{4} (|\tilde{z}| + |\tilde{w}|).
\end{aligned}$$

Since  $|\tilde{z}| + |\tilde{w}| < \delta_1$ , another application of regular differentiability in (8.7) implies

$$|f_{x_i, T_i}(\tilde{z}, \tilde{w}) - f_{x_i, T_i}(\tilde{z}, 0) - L_i \tilde{w}| \leq \xi(|\tilde{z}| + |\tilde{w}|) \leq \frac{s_i}{4} (|\tilde{z}| + |\tilde{w}|).$$

Hence we obtain, by triangle inequality and the original choice of  $\tilde{z}$  and  $\tilde{w}$ ,

$$\left( \|f_{x_i, T_i} - f_{\varphi(u), T_i}\|_r - \frac{\tau}{2} \right) (|\tilde{z}| + |\tilde{w}|) \leq s_i (|\tilde{z}| + |\tilde{w}|)$$

which implies  $\Delta_i((x_i, T_i), (\varphi(u), \varphi'_{n-1}(u))) \leq \tau/2$  so  $u \notin P$ .

We have established (9.3) and now show (9.2). We first estimate, using the fact  $\alpha$  is small and  $\tilde{w} \in \mathbb{R}^{n-1}$ ,

$$\begin{aligned} & |\varphi(u + \tilde{w}) - \varphi(u) - T_i \tilde{w}| \\ & \leq (\|T_i - T_\infty\| + \|T_\infty - \tilde{T}_\infty\| + \text{Lip}(\gamma)) |\tilde{w}| \\ & \leq \left( \frac{s_i}{8\text{Lip}(f)} + \kappa + 2\kappa \right) |\tilde{w}| \\ & \leq \frac{s_i}{4\text{Lip}(f)} |\tilde{w}|. \end{aligned}$$

Combining this with (9.3),

$$\begin{aligned} & (|\tilde{u}| + |\tilde{z}|) \text{reg}_{n-1} g(u, B(u, s)) \geq |g(u + \tilde{u}) - g(u + \tilde{z})| \\ & = |f(\varphi((u + \tilde{z} + \tilde{w}))) - f(\varphi((u + \tilde{z}))) - L_\infty(\tilde{w})| \\ & \geq |f(\varphi(u + \tilde{w}) + \tilde{z}) - f(\varphi(u) + \tilde{z}) - L_\infty(\tilde{w})| \\ & \quad - |f(\varphi(u + \tilde{w}) + \tilde{z}) - f(\varphi(u) + \tilde{z}) \\ & \quad \quad - f(\varphi((u + \tilde{z} + \tilde{w}))) + f(\varphi((u + \tilde{z})))| \\ & \geq |f_{\varphi(u), T_i}(\tilde{z}, \tilde{w}) - f_{\varphi(u), T_i}(\tilde{z}, 0) - L_\infty(\tilde{w})| \\ & \quad - \text{Lip}(f) |\varphi(u + \tilde{w}) - \varphi(u) - T_i \tilde{w}| \\ & \quad - 2\text{Lip}(f) \text{Lip}(\varphi) |\tilde{w}| \\ & \geq |f_{\varphi(u), T_i}(\tilde{z}, \tilde{w}) - f_{\varphi(u), T_i}(\tilde{z}, 0) + f_{x_i, T_i}(\tilde{z}, 0) - f_{x_i, T_i}(\tilde{z}, \tilde{w})| \\ & \quad - |f_{x_i, T_i}(\tilde{z}, \tilde{w}) - f_{x_i, T_i}(\tilde{z}, 0) - L_i(\tilde{w})| - \|L_i - L_\infty\| |\tilde{w}| \\ & \quad - \frac{s_i}{4} |\tilde{w}| - 6\text{Lip}(f) \kappa |\tilde{w}| \\ & \geq \left( \|f_{x_i, T_i} - f_{\varphi(u), T_i}\|_r - \frac{\tau}{2} \right) (|\tilde{z}| + |\tilde{w}|) \\ & \quad - |f_{x_i, T_i}(\tilde{z}, \tilde{w}) - f_{x_i, T_i}(\tilde{z}, 0) - L_i(\tilde{w})| - \frac{s_i}{8} |\tilde{w}| \\ & \quad - \frac{s_i}{4} |\tilde{w}| - \frac{s_i}{8} |\tilde{w}|. \end{aligned}$$

For the first term we use the fact  $u \in P$  implies  $\|f_{x_i, T_i} - f_{\varphi(u), T_i}\|_r > \tau + s_i$ . For the second term we use regular differentiability (8.7). Hence the above expression is greater or equal than

$$\begin{aligned} & s_i(|\tilde{z}| + |\tilde{w}|) - \xi(|\tilde{z}| + |\tilde{w}|) - \frac{s_i}{2} |\tilde{w}| \geq \frac{s_i}{4} (|\tilde{z}| + |\tilde{w}|) \\ & \geq \frac{s_i}{8} (|\tilde{u}| + |\tilde{z}|) \end{aligned}$$

which proves (9.2).

Let  $Q_0 = \{u \in Q : B(u, s) \subset Q\}$ . Then, from Proposition 3.3,

$$\begin{aligned} & (\mathcal{L}^{n-1} \times \mu)(P \cap Q_0) \\ & \leq \frac{C(\mu)(5\text{Lip}(f))^{2(n+C(\mu))-4}}{(s_i/8)^{2(n+C(\mu))-2}} \int_Q \|g'_{n-1}\|^2 d(\mathcal{L}^{n-1} \times \mu). \end{aligned}$$

Recall  $\Omega$  is a cylinder of height  $r$  whose cross sections are balls in  $\mathbb{R}^{n-2}$  of radius  $r$ ; hence  $\mathcal{L}^{n-1}(\Omega) = \omega_{n-2} r^{n-1}$ . We estimate, using the Bernoulli inequality and the choice of  $s$  and  $\xi$ ,

$$\begin{aligned}
\frac{(\mathcal{L}^{n-1} \times \mu)(Q_0)}{(\mathcal{L}^{n-1} \times \mu)(Q)} &\geq \left(1 - \frac{2s}{r}\right) \left(1 - \frac{s}{r}\right)^{n-2} \frac{\mu(B(0, tr - s))}{\mu(B(0, tr))} \\
&\geq \left(1 - \frac{ns}{r}\right) \frac{\mu(B(0, tr - s))}{\mu(B(0, tr))} \\
&\geq (1 - \tau).
\end{aligned}$$

Using the fact  $\Delta_i((x_i, T_i), (\varphi, \varphi'_{n-1})) \leq 1$  and our choice of  $\sigma_i$  we conclude,

$$\begin{aligned}
&\sigma_i \int_Q (\Delta_i((x_i, T_i), (\varphi, \varphi'_{n-1})) - \Delta_i((x_i, T_i), (x_\infty, T_\infty))) \, d(\mathcal{L}^{n-1} \times \mu) \\
&\leq \sigma_i \int_P \Delta_i((x_i, T_i), (\varphi, \varphi'_{n-1})) \, d(\mathcal{L}^{n-1} \times \mu) + \sigma_i \tau (\mathcal{L}^{n-1} \times \mu)(Q \setminus P) \\
&\leq \sigma_i (\mathcal{L}^{n-1} \times \mu)(P \cap Q_0) + \sigma_i (\mathcal{L}^{n-1} \times \mu)(Q \setminus Q_0) + \sigma_i \tau (\mathcal{L}^{n-1} \times \mu)(Q) \\
&\leq \sigma_i \frac{C(\mu)(5\text{Lip}(f))^{2(n+C(\mu))-4}}{(s_i/8)^{2(n+C(\mu))-2}} \int_Q \|g'_{n-1}\|^2 \, d(\mathcal{L}^{n-1} \times \mu) \\
&\quad + 2\sigma_i \tau (\mathcal{L}^{n-1} \times \mu)(Q) \\
&\leq 2^{-i-3} \left( \int_Q \|g'_{n-1}\|^2 \, d(\mathcal{L}^{n-1} \times \mu) + \tau (\mathcal{L}^{n-1} \times \mu)(Q) \right).
\end{aligned}$$

For  $i > K$ , we note that  $i > K$  implies  $\varepsilon > 1/i$  and the choice of  $\sigma_i, \kappa$  and  $\tau$  give,

$$\sigma_i \leq \frac{2^{-i-5}}{i+1} \frac{s_i}{48\text{Lip}(f)} \leq 2^{-i-3} \frac{\varepsilon \kappa}{4} = 2^{-i-3} \tau.$$

Hence we obtain, for  $i > K$ ,

$$\begin{aligned}
&\sigma_i \int_Q (\Delta_i((x_i, T_i), (\varphi, \varphi'_{n-1})) - \Delta_i((x_i, T_i), (x_\infty, T_\infty))) \, d(\mathcal{L}^{n-1} \times \mu) \\
&\leq 2^{-i-3} \tau (\mathcal{L}^{n-1} \times \mu)(Q).
\end{aligned}$$

Adding the inequalities together, for different  $i$ , we obtain,

$$\int_Q \Delta(\varphi) \, d(\mathcal{L}^{n-1} \times \mu) \leq \frac{1}{4} \int_Q \|g'_{n-1}\|^2 \, d(\mathcal{L}^{n-1} \times \mu) + \frac{\tau}{4}.$$

**9.5. Combining the Estimates.** It remains to analyze the value of the first term in (9.1). We use the fact  $f'(\varphi; \varphi'_{n-1}) = g'_{n-1} + L_\infty$  to obtain,

$$\begin{aligned}
&\int_Q (\|f'(\varphi; \varphi'_{n-1})\|_H^2 - \|L_\infty\|_H^2) \, d(\mathcal{L}^{n-1} \times \mu) \\
&= \int_Q \|g'_{n-1}\|_H^2 \, d(\mathcal{L}^{n-1} \times \mu) + \int_Q \langle 2L_\infty, g'_{n-1} \rangle_H \, d(\mathcal{L}^{n-1} \times \mu).
\end{aligned}$$

By putting together the previous estimates of this section and using the definition of  $R_\infty$  we find (all integrals are with respect to  $\mathcal{L}^{n-1} \times \mu$ ),

$$\begin{aligned}
& \int_Q (h_\infty(x_\infty, T_\infty) - h_\infty(\varphi, \varphi'_{n-1})) \\
& \geq \int_Q \|g'_{n-1}\|_H^2 + \int_Q \langle 2L_\infty, g'_{n-1} \rangle_H - \frac{1}{4} \int_Q \|g'_{n-1}\|_H^2 \\
& \quad + \int_Q \left\langle 2 \sum_{i=0}^{\infty} \gamma_i (L_i - L_\infty), g'_{n-1} \right\rangle_H - \frac{1}{4} \int_Q \|g'_{n-1}\|_H^2 - \frac{9\tau}{4} \\
& \geq \int_Q \langle R_\infty, g'_{n-1} \rangle - \frac{9\tau}{4} \\
& = \langle R_\infty e, w \rangle + \int_Q \langle R_\infty, g'_{n-1} - w \otimes e \rangle_H - \frac{9\tau}{4} \\
& \geq \frac{7\tau}{4} - \|R_\infty\|_H \left\| \int_Q (g'_{n-1} - w \otimes e) \right\|_H.
\end{aligned}$$

Hence to conclude our proof it suffices to show that

$$\|R_\infty\|_H \left\| \int_Q (g'_{n-1} - w \otimes e) \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H < \frac{7\tau}{4}.$$

**9.6. Estimate of  $\left\| \int_Q (g'_{n-1} - w \otimes e) \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H$ .** We define the maps  $\zeta, \tilde{\zeta}: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$  by

$$\zeta(u) = g(\langle u, e \rangle e) \text{ and } \tilde{\zeta}(u) = \zeta(u) - g(0) - \langle u, e \rangle w.$$

By Lemma 8.7 we have, for every  $u \in \overline{Q}$ ,

$$|g(u) - \zeta(u)| \leq \frac{\tau r}{n(1 + \|R_\infty\|_H)}.$$

We now use the following inequality [14, Corollary 9.4.2] which arises from an application of the divergence theorem.

**Lemma 9.1.** *Let  $\Pi \subset \mathbb{R}^{n-1}$  be a bounded open set with Lipschitz boundary. Then for every Lipschitz function  $\Lambda: \overline{\Pi} \rightarrow \mathbb{R}^m$ ,*

$$\left\| \int_\Pi \Lambda'(u) \, d\mathcal{L}^{n-1}(u) \right\|_H \leq \mathcal{H}^{n-2}(\partial\Pi) \max_{u \in \partial\Pi} |\Lambda(u)|.$$

For each  $b \in (-tr, tr)$  we apply the lemma to the map  $z \mapsto g(z, b) - \zeta(z, b)$  for  $z \in \Omega$ . Recall  $\Omega \subset \mathbb{R}^{n-1}$  is a cylinder of height  $r$  whose cross sections are balls in  $\mathbb{R}^{n-2}$  of radius  $r$ ; hence  $\mathcal{H}^{n-2}(\partial\Omega) = 2\omega_{n-2}r^{n-2} + (n-2)\omega_{n-2}r^{n-2}$ .

We estimate,

$$\begin{aligned}
& \left\| \int_Q (g'_{n-1} - \zeta'_{n-1}) \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H \\
& \leq \int_{(-tr, tr)} \left\| \int_{\Omega} (g'_{n-1} - \zeta'_{n-1}) \, d\mathcal{L}^{n-1} \right\|_H \, d\mu \\
& \leq \int_{(-tr, tr)} \frac{\tau r}{n(1 + \|R_{\infty}\|_H)} \mathcal{H}^{n-2}(\partial\Omega) \, d\mu \\
& = \frac{\mu(-tr, tr) \tau r (2\omega_{n-2} r^{n-2} + (n-2)\omega_{n-2} r^{n-2})}{n(1 + \|R_{\infty}\|_H)} \\
& = \frac{\tau(\mathcal{L}^{n-1} \times \mu)(Q)}{(1 + \|R_{\infty}\|_H)}.
\end{aligned}$$

Note that the function  $\tilde{\zeta}$  depends only on the projection onto  $\mathbb{R}e$ . Define  $\Lambda: \mathbb{R} \rightarrow \mathbb{R}^{n-1}$  by  $\Lambda(\theta) = \tilde{\zeta}(\theta e)$  so that  $\tilde{\zeta}(u) = \Lambda(\langle u, e \rangle)$ . It follows from Rademacher's theorem that  $\Lambda'(\theta)$  exists for almost every  $\theta$ . Hence, using Fubini's theorem and the chain rule, we see  $\tilde{\zeta}'_{n-1}(u) = \Lambda'(\langle u, e \rangle) \otimes e$  for  $\mathcal{L}^{n-1} \times \mu$  almost every  $u$ . Let  $S_0 = \{u \in \overline{\Omega} : \langle u, e \rangle = 0\}$ . Then, by Fubini's theorem,

$$\begin{aligned}
\left\| \int_Q \tilde{\zeta}'_{n-1} \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H & \leq \int_{B(0, tr)} \left\| \int_{\Omega} \Lambda'(\langle u, e \rangle) \otimes e \, d\mathcal{L}^{n-1} \right\|_H \, d\mu \\
& \leq \mu(B(0, tr)) \left\| \int_{S_0} (\Lambda(0) - \Lambda(-r)) \otimes e \, d\mathcal{H}^{n-2} \right\|_H \\
& = |\tilde{\zeta}(-re)| \mathcal{H}^{n-2}(S_0) \mu(B(0, tr)) \\
& = |g(0) - g(-re) - rw| \mathcal{H}^{n-2}(S_0) \mu(B(0, tr)).
\end{aligned}$$

By using the expressions for  $g$  and  $w$  one can check,

$$\begin{aligned}
& g(0) - g(-re) - rw \\
& = -[f(x_{\infty} + T_{\infty}(-re)) - f(x_{\infty}) - L_{\infty}(-re)] \\
& \quad + [f(x_{\infty} - T_{\infty}re) - f(x_{\infty} - \tilde{T}_{\infty}re + \tilde{x} - \kappa r \langle e, \tilde{e} \rangle \tilde{x}/|x|)] \\
& \quad + [f(x_{\infty} + \tilde{x}) - f(x_{\infty} + x)].
\end{aligned}$$

For the first term, as  $|r| < \delta_1$ , we may use (8.7) to bound it by  $\xi r$ . Since  $\kappa r = |x|$ , the other terms can be made small, relative to  $\xi r$ , by choosing  $\alpha$  small. Hence we obtain, by choice of  $\xi$ ,

$$\begin{aligned}
\left\| \int_Q \tilde{\zeta}'_{n-1} \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H & \leq 2\xi r \mathcal{H}^{n-2}(S_0) \mu(B(0, tr)) \\
& \leq 2\xi \mathcal{L}^{n-1}(\Omega) \mu(B(0, tr)) \\
& \leq \frac{\tau(\mathcal{L}^{n-1} \times \mu)(Q)}{2(1 + \|R_{\infty}\|_H)}.
\end{aligned}$$

Note that  $\tilde{\zeta}'_{n-1} = \zeta'_{n-1} - w \otimes e$ . Hence we can estimate,

$$\begin{aligned} & \|R_\infty\|_H \left\| \int_Q (g'_{n-1} - w \otimes e) \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H \\ & \leq \|R_\infty\|_H \left\| \int_Q (g'_{n-1} - \zeta'_{n-1}) \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H \\ & \quad + \|R_\infty\|_H \left\| \int_Q (\zeta'_{n-1} - w \otimes e) \, d(\mathcal{L}^{n-1} \times \mu) \right\|_H \\ & \leq \tau + \frac{\tau}{2} < \frac{7\tau}{4}. \end{aligned}$$

This completes the proof of (8.8) which proves Proposition 8.4 and hence Theorem 1.1.

*Remark 9.2.* When considering mappings from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ ,  $m < n$ , we may replace in the above arguments the measure  $\mathcal{L}^{n-1} \times \mu$  with  $\mathcal{L}^m \times \nu$  where  $\nu$  is a doubling measure on  $\mathbb{R}^{n-m}$ . By Wu [18], such  $\nu$  can be chosen of arbitrarily small Hausdorff dimension. Together with the easy fact that any set is contained in a  $G_\delta$  set of the same Hausdorff dimension, this leads to a set of Hausdorff dimension arbitrarily close to  $m$  that contains points of differentiability of every Lipschitz function  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

#### REFERENCES

- [1] G. Alberti, M. Csörnyei, and D. Preiss. Differentiability of Lipschitz Functions, Structure of Null Sets, and Other Problems. *Proceedings of the International Congress of Mathematicians*, III:1379–1394, 2010.
- [2] David Bate. Structure of Measures in Lipschitz Differentiability Spaces. arXiv:1208.1954.
- [3] Y. Benyamini and J. Lindenstrauss. *Geometric Nonlinear Functional Analysis, Volume 1*, volume 48. American Mathematical Society Colloquium Publications, 2000. ISBN 0-8218-0835-4.
- [4] J. Cheeger. Differentiability of Lipschitz Functions on Metric Measure Spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.
- [5] M. Csörnyei and P. Jones. Product Formulas for Measures and Applications to Analysis and Geometry. URL: [www.math.sunysb.edu/Videos/dfest/PDFs/38-Jones.pdf](http://www.math.sunysb.edu/Videos/dfest/PDFs/38-Jones.pdf).
- [6] T. De Pauw and P. Huovinen. Points of  $\varepsilon$ -Differentiability of Lipschitz Functions From  $\mathbb{R}^n$  to  $\mathbb{R}^{n-1}$ . *Bull. London Math. Soc.*, 34(5):539–550, 2002.
- [7] J. Dugundji. *Topology*. Allyn and Bacon, Inc., Boston, Mass., 1966. ISBN 978-0205002719.
- [8] T. Fowler and D. Preiss. A Simple Proof of Zahorski’s Description of Non-differentiability Sets of Lipschitz Functions. *Real Anal. Exchange*, 34(1):127–138, 2009.
- [9] J. Heinonen. *Lectures on Analysis on Metric Spaces*. Universitext. Springer-Verlag, New York, 2001. ISBN 0-387-95104-0.
- [10] A. S. Kechris. *Classical Descriptive Set Theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. ISBN 0-387-94374-9.



- [11] S. Keith. A Differentiable Structure for Metric Measure Spaces. *Adv. Math.*, 183(2):271–315, 2004.
- [12] J. Lindenstrauss and D. Preiss. Almost Fréchet Differentiability of Finitely Many Lipschitz Functions. *Mathematika*, 43(2):393–412, 1996.
- [13] J. Lindenstrauss and D. Preiss. On Fréchet Differentiability of Lipschitz Maps Between Banach Spaces. *Ann. of Math. (2)*, 157(1):257–288, 2003.
- [14] J. Lindenstrauss, D. Preiss, and J. Tišer. *Fréchet Differentiability of Lipschitz Functions and Porous Sets in Banach Spaces*, volume 179 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2012. ISBN 978-0-691-15356-8.
- [15] D. Preiss. Differentiability of Lipschitz Functions on Banach Spaces. *J. Funct. Anal.*, 91(2):312–345, 1990.
- [16] G. Speight. Surfaces Meeting Porous Sets in Positive Measure. *Israel Journal of Mathematics*, 2013. doi: 10.1007/s11856-012-0171-4.
- [17] P. Tukia. Hausdorff Dimension and Quasisymmetric Mappings. *Math. Scand.*, 65(1):152–160, 1989.
- [18] J. Wu. Hausdorff Dimension and Doubling Measures on Metric Spaces. *Proc. Amer. Math. Soc.*, 126(5):1453–1459, 1998.
- [19] Z. Zahorski. Sur L’Ensemble des Points de Non-Derivabilité d’une Fonction Continue. *Bull. Soc. Math. France*, 74:147–178, 1946.
- [20] L. Zajíček. Porosity and  $\sigma$ -Porosity. *Real Anal. Exchange*, 13(2):314–350, 1987/88.
- [21] L. Zajíček. On  $\sigma$ -Porous Sets in Abstract Spaces. *Abstr. Appl. Anal.*, 2005(5):509–534, 2005.